Compensation Methods in a Competitive Labor Market: The Role of Asymmetric Information

Felipe Balmaceda[1]

Department of Economics, Ilades-Georgetown University[2]

May 30, 2000

JEL classification: J31, J33, D82.

Keywords: Incentive pay, straight salaries, screening, monitoring.

[1]I would like to thank John Abowd, David Easley, Dave Kaplan, Rob Olsen, Michael Waldman and seminar participants at Universidad Carlos III-Madrid, CERGE-EI, Universidas de Chile, and Cornell University Labor Workshop for their useful commnets to an earlier version. As usual, all remaining errors are of my own responsibility.

[2]Address: Erasmo Escala 1835, Rm. 263, Zip Code: 6500620, Santiago-Chile, E-mail: fbalmace@uahurtado.cl.

Abstract

In this paper we develop an asymmetric information model that provides a rationale for the existence of pay-for-performance contracts in the absence of incentive for effort and explains when and in which occupations pay-for-performance is more likely to be observed. In our model competition among firms for the best workers forces firm to link pay to performance in order to provide the best workers with a higher expected compensation. Furthermore, the model predicts among other things and contrary to the moral hazard model, that there is an equilibrium in which workers under contracts with a larger pay-for-performance sensitivity exert less effort than workers under contracts with a smaller pay-for-performance sensitivity.

The paper also makes contributions to the theoretical literature on screening games. It is shown that in a competitive market and under a slightly modified timing than the one proposed by Rothschild and Stiglitz' (1976) a unique equilibrium exits when a appropiately chosen equilibrium refinement is used and that the standard result in screening games in monopolistic settings known as no distortion at the top (see, Laffont and Tirole, 1996) does not hold in a competitive market.

1 Introduction

Evidence shows that some workers are compensated according to pay-for-performance and some according to straight salaries. In the US, for instance, between 20% and 30% of the workforce is covered by some kind of pay-for-performance compensation method. The standard rationale given for the use of pay-for-performance contracts is that this compensation method provides agents with incentives to work harder relative to the case in which they are paid straight salaries. While appealing, this explanation is not always easily reconciled with theory and empirical evidence. Prendergast (1998, 1999) concludes that the selection effects of pay-for-performance contracts are roughly of equal size to the incentives effects, despite the overwhelming focus of the incentive effect on the literature. Thus, one can conclude that to focus on the incentive effect of pay-for-performance results in a lack of understanding of the role played by incentive contracts and bias the empirical work that tries to identify the effects of contracts on productivity and compensation.[1]

In this paper rather than emphasizing the incentive effect of pay-for-performance, the sorting or selection effect of pay-for-performance contracts in a competitive labor market is stressed.[2] In so doing, a simple asymmetric information model is developed in which risk neutral firms compete for risk averse workers of different abilities, which is unobserved by employers. Within this framework, we show that when the proportion of high-ability workers is small, pay-for-performance and straight-salary contracts are simultaneously observed, while for a large proportion of high-ability workers, only pay-for-performance contracts are observed.[3] This results in that within an occupation, pay-for-performance workers gain more

[1] See, for instance, Lazear (1999), Paarsh and Shearer (1997) and Shearer (1997).

[2] Maskin and Riley (1985) and Khalil and Lawaree (1995) study somewhat similar issues in monopolistic settings. The former study the efficiency of output and input monitoring in terms of tax revenue and the latter study how residual claimancy affects the choice between these two types of monitoring.

[3] Lazear (1986) presents a similar result to ours based also in a sorting rationale. Hermalin (1995) in a principal-agent setting, also shows that identical firms can offer different incentive contracts to their

and are more productive than straight-salary workers. Furthermore, the model predicts that when only pay-for-performance contracts are offered, the pay-for-performance sensitivity is smaller than when pay-for-performance and straight-salary contracts are simultaneously offered. That is, the model predicts not only which types of contracts should be observed within an occupation, but also that the form of the observed pay-for-performance contracts differ depending on whether these are observed together with straight-salary contracts or by themselves only. Thus, the results presented in this paper coupled with the evidence, for instance, summarized in Prendergast (1998, 1999) suggests that too much attention to the incentive hypothesis has crowded out the importance of sorting in explaining not only the heterogeneity in compensation, but also the form of compensation contracts.

The model, however, has the awkward prediction that in no equilibrium only straight-salary contracts are observed and, furthermore, effort is ignored. Thus, we extend our model to deal with monitoring and effort in the simplest way possible. When monitoring is introduced, it is shown that for jobs in which the monitoring difficulty is not too large our results hold, while when the monitoring difficulty is large, only straight-salary contracts are observed. Furthermore, it is shown that only those occupations in which monitoring takes place use pay-for-performance and pay more than in occupations in which no monitoring takes place. When effort is introduced, we also show that our results are robust to effort and that under certain conditions, workers under straight salaries exert more effort and have a larger expected productivity than workers under pay-for-performance. Thus, it is not always the case that pay-for-performance workers produce and earn more than straight-salary workers. In addition, it is shown that a performance standard and the pay-for-performance sensitivity are substitutes devices to achieve perfect sorting. When workers are highly risk averse, a high performance standard coupled with a small pay-for-performance sensitivity is used to sort workers out, while when workers are less risk averse, a low performance standard

managers. Yet, the rationale for his results is completely different from the one provided here.

coupled with a large pay-for-performance sensitivity is used to achieve perfect sorting.

Finally, the paper also contributes to the theory of screening in competitive markets by showing that when firms can reject applicants and an appropiated equilibrium refinement is adopted a unique equilibrium *always* exists and this entails pooling in the region in which Rothschild and Stiglitz's (1974) insurance model has no equilibrium.[4] For instance, our equilibrium concept can be used to show that debt contracts that include a collateral may exist for other reason than sorting different risks out and to avoid rationing (see, Bester, 1985) and to show that under certain conditions pool risks is welfare enhancing.

The outline for the rest of the paper is as follows. Section 2 describes a pure screening model in which output depends only on a worker's ability. Section 3 presents a benchmark case and derives the unique Perfect Bayesian equilibrium. In section 4, we pursue two extensions. We extend the model to deal with effort first and then endogenous and imperfect monitoring. In the next section, section 5, the predictions of the model are contrasted with empirical evidence and predictions of similar models. Finally, concluding remarks are presented in the last section.

2 The Basic Model and The Equilibrium Concept

2.1 The Basic Model

Identical risk neutral firms compete for a fixed number of workers, where labor supply is fixed at one unit for each individual and the price of output is normalized to 1. Employees differ in their innate ability, θ, and each one knows his own ability. The employers know only that a worker's ability takes on two values, $\theta \in \{\bar{\theta}, \underline{\theta}\}$, where $1 > \bar{\theta} > \underline{\theta} \geq 0$, with prior probability $\mu \equiv pr(\theta = \overline{\theta})$. Here, μ is interpreted as the population share of high-ability workers.

[4] Under the standard timing, *if an equilibrium exists*, it always entails separation.

Each firm i has the same production technology; that is, there are two possible output levels $\bar{y}$ and $\underline{y}$, with $\triangle y$ defined by $\bar{y} - \underline{y}$ greater than zero. It is assumed that a θ-worker's probability of producing the high output $\bar{y}$ is θ (probability of success, hereafter). Therefore, a θ-worker's expected output, denoted by $y(\theta)$, is $\theta\bar{y}+(1-\theta)\underline{y}$, and since, $\bar{\theta} > \underline{\theta}$, $y(\bar{\theta}) > y(\underline{\theta})$.

Each firm i offers a menu of contracts denoted by C, where a contract $C_k \in C$ specifies a non-negative wage paid for each output level: $\overline{w}_k$ for $\bar{y}$ and $\underline{w}_k$ for $\underline{y}$, with $\overline{w}_k \geq \underline{w}_k$. If $\overline{w}_k$ is larger than $\underline{w}_k$, the contract is a pay-for-performance contract, otherwise it is a straight salary contract. Furthermore, unless otherwise noted, zero monitoring costs are assumed; that is, firms can verify output at no cost, and workers' reservation utility is assumed to be zero. The subscript θ will be omitted where obvious. A worker's utility from compensation w is $U(w)$, where $U^{'}(w) > 0, U^{''}(w) < 0$. Hence, a θ-worker's expected utility when he accepts contract C_k is given by,[5]

$$V(C_k \mid \theta) = \theta U(\overline{w}_k) + (1-\theta)U(\underline{w}_k). \tag{1}$$

Similarly, a firm's expected profit from employing a θ-worker under the contract C_k is given by,

$$\Pi(C_k \mid \theta) = \theta(\bar{y} - \overline{w}_k) + (1-\theta)(\underline{y} - \underline{w}_k). \tag{2}$$

The timing of decisions adopted here was suggested by Hellwig (1987) and it is as follows. At *Stage 1*, firms are symmetrically informed and simultaneously offer a menu of contracts that includes either a pay-for-performance or straight salary contract or both for the up-coming period. At *Stage 2,* after offers have been made, each worker applies to a particular

[5] Note that, in the space of performance-contingent wages, workers' indifference curves satisfy the Spence-Mirrlees's single-crossing property. In our setting, this means that the extra wage that must be paid for good performance to compensate a worker for the risk that a pay-for-performance contract imposes is lower for high-ability workers.

firm for the upcoming period. In the case that more than one firm offers the same contract, workers choose randomly between firms. At the third stage, *Stage 3,* after each worker has chosen a contract and *firms have observed other firms' offers*, firms have the opportunity to either accept or reject a worker's application. Yet, once a worker has agreed to work for a particular firm and has been accepted, the terms of the agreement become binding for that period. Finally, output is realized and compensation takes place as specified in the contract.

2.2 The Equilibrium Concept

This section briefly explains the equilibrium concept that will be used and the importance of the timing adopted. The reason being that under the standard equilibrium concepts and the standard timing for screening games[6] this type of model generally have problems ranging from *non-existence of equilibrium* for some parameters value to *multiple equilibria* for some others.

Hellwig (1987) added the third stage to the two-stage screening game in order to solve the competitive screening games' known *non-existence* of equilibrium problem. Because the last two stages mimic a signaling game, however, Hellwig's timing effectively trades the problem of non-existence for the problem of multiple equilibrium. In particular, as has been known since Cho and Kreps (1987), signaling games have a plethora of Perfect Bayesian equilibrium (hereafter, PBE) that are supported by unreasonable off-the-equilibrium path beliefs. In this paper, we adopt a signaling equilibrium refinement proposed by Mailath et al. (1993) to eliminate equilibria that are based on unreasonable beliefs. In particular, we will require that any PBE of the signaling sub-game (stages 2 and 3) must be *undefeated* among all possible PBEs that can arise from any first stage contract offers. In our model, this refinement places the following restrictions on the off-the-equilibrium path belief function.

[6]The classic example of this is Rothschild's and Stiglitz's (1976) competitive insurance model, which includes only our stage 1 and 2.

Consider a proposed PBE and a given contract C_k that is not chosen in this equilibrium, but it is chosen by some types of workers in an alternative PBE. Let K be the set of workers' types that choose C_k in the alternative equilibrium. If each member of K prefers the alternative equilibrium to the proposed one, with strict preference for at least one type, then after observing C_k the firms' belief that a worker is of high-ability given that C is offered must satisfy

$$\mu\left(\overline{\theta} \mid C_k\right) = \frac{\mu(\overline{\theta})\beta(\overline{\theta})}{\mu(\overline{\theta})\beta(\overline{\theta}) + (1 - \mu(\underline{\theta}))\beta(\underline{\theta})},$$

with $\beta(\theta) = 0$ for all $\theta \notin K$, and $\beta(\theta) = 1$ for all $\theta \in K'$, where $K' \subset K$ is the set of workers' types who strictly prefer the alternative PBE to the proposed one.[7]

Intuitively, a proposed PBE is said to be defeated by an alternative PBE if there are deviations from the proposed PBE which are played in the alternative PBE by some workers types, all of whom prefer the alternative equilibrium to the proposed one.

We adopt the undefeated equilibrium refinement over the Intuitive Criterion and others because the equilibrium selected by the later remains unchanged for any positive proportion of low-ability workers, the situation is quite different when there is no low-ability workers. Then contract choice serves no separating purpose and the only equilibrium of interest is the one in which the workers (who are all high-ability) are all paid a straight salary equal to their productivity. It seems unreasonable that the outcome of a game with 1 worker in a million chance of a low-ability worker differs significantly from a game in which there is no chance of such a worker. It is reasonable to think that we are not certain about the distributions of type, therefore, the model and the equilibrium of our model can be useful only if the predicted outcome is not overly sensitive to the description of the environment.

[7] The appendix contains a formal definition of Undefeated Perfect Bayesian Equilibrium.

3 The Analysis

3.1 The Benchmark

For the reader interested in the formal analysis and the game-theoretic issues involved, the appendices provide formal statements of all propositions and, unless otherwise noted, all proofs are in the appendices.

Equations 1 and 2 tell us that under complete information, all contracts should be straight salary contracts. If $\overline{w}_k$ were different from $\underline{w}_k$, then a reduction in $\overline{w}_k$ with the appropriate increase in $\underline{w}_k$ would shift some of the risk of the contract from a worker to an employer, which is advantageous because employers are risk neutral, and workers are not. Therefore, in equilibrium, the contract offered to both, high- and low-ability workers, is a straight salary contract, and competition among employers forces firms to pay each type her expected output. However, equations 1 and 2 also suggest that under incomplete information, the difference between $\overline{w}_k$ and $\underline{w}_k$ may serve as a sorting device. Because high-ability workers have a higher probability of success an increase in $\overline{w}_k$ in return for a reduction in $\underline{w}_k$ lowers their expected payoff less than it lowers low-ability workers' expected payoff. This intuition is explored in the following sub-section.

3.2 The Incomplete Information Case

As suggested above, under incomplete information firms may offer an optimally chosen pay-for-performance contract to take advantage of high-ability workers' willingness to accept an increase in $\overline{w}$ in return for a reduction in $\underline{w}$ to reveal their ability. In this section it is shown that for certain parameter values there is a unique equilibrium in which some firms use pay-for-performance and others straight salaries to sort workers out, while for different parameter values there is a unique equilibrium in which all firms offer the same pay-for-performance contract and workers are not separated out.

In the separating equilibrium high-ability workers work for a pay-for-performance firm and low-ability workers work for a straight salary firm. Since perfect sorting takes place, competition forces employers to pay each ability type her expected output and, therefore, high-ability workers' expected compensation is strictly higher than low-ability workers' expected compensation. For high-ability workers, this perfect sorting has the advantage that they do not subsidize low-ability workers. The disadvantage, however, is that pay-for-performance contracts imposes risk, which under complete information the employers would bear.

In the pooling equilibrium all firms offer the same pay-for-performance contract to high- and low-ability workers. Since sorting does not take place high-ability workers subsidize low-ability workers and competition forces firms to offer the contract that high-ability workers prefer the most among all potential contracts that break even on the population average probability of success, $\widehat{\theta} = \mu\overline{\theta} + (1 - \mu)\underline{\theta}$. There are two things to note here. First, even if both types are offered the same contract, high-ability workers have a higher expected compensation. The reason being that they have a larger probability of success and, therefore, are more likely to benefit from the higher compensation that a pay-for-performance contract attaches to a high output realization. Second, the risk imposed by the pay-for-performance contract in the pooling equilibrium is lower than the one imposed by the pay-for-performance contract in the separating equilibrium. The reason being that in the separating equilibrium enough risk must be imposed to discourage low-ability workers from mimicking high-ability workers, while in the pooling equilibrium only a small amount of risk is imposed to reduce the subsidy from high-ability workers to low-ability workers.[8]

[8] To better understand the nature of the pooling contract we need first to understand the cost of pooling all workers under a straight-salary, w. Note, first, that competition implies that w must be such that firms break even at the population average probability of success, $\widehat{\theta}$, and at w, both, high- and low-ability workers' expected wage is the same. This translates into high-ability workers giving a subsidy to low-ability workers equivalent to the difference between w and low-ability workers' expected output. What if firms offer a

In the pooling equilibrium the main cost for high-ability workers is that they subsidize low-ability workers, while in the separating equilibrium the main cost is that they face too much risk. *This cost of sorting will exceed the benefit* if the population share, μ, of high-ability workers is large enough so that the population average probability of success, $\widehat{\theta}$, is close enough to $\overline{\theta}$ because then the cost of subsidizing low-ability workers is small. If this is the case, both, high- and low-ability workers obtain a larger expected payoff when sorting is not attempted. Hence, when $\widehat{\theta}$ is close enough to $\overline{\theta}$, *i.e.*, when the population share of high-ability workers is larger than a threshold $\tilde{\mu}$, in equilibrium competition forces firms to offer the pay-for-performance contract that high-ability workers prefer the most among all those contracts that break even on the population average, and this contract is accepted by all workers.

Formally, in a separating equilibrium firms maximize high-ability workers' expected payoff subject to: high- and low-ability workers' incentive compatibility constraints, $(IC\overline{\theta})$ and $(IC\underline{\theta})$, respectively; low-ability workers' participation constraint, $(IR\underline{\theta})$; and a zero expected profit constraint in each contract offered, $(ZP\overline{\theta})$ and $(ZP\underline{\theta})$, respectively. We will define the contracts $C^s_{\overline{\theta}} \equiv (\overline{w}^s_{\overline{\theta}}, \underline{w}^s_{\overline{\theta}})$ and $C^s_{\underline{\theta}} \equiv (\overline{w}^s_{\underline{\theta}}, \underline{w}^s_{\underline{\theta}})$ as the solution to the following program,

$$\max_{\{\overline{w}^s_\theta, \underline{w}^s_\theta\}^{\overline{\theta}}_{\underline{\theta}}} \overline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\overline{\theta}}) \qquad (PS)$$

subject to

contract with a slight amount of risk? Low-ability workers' expected wage would be smaller than w because they have a lower probability of success than the average worker, while high-ability workers' expected wage would be larger because they have a larger probability of success than the average worker. Therefore, the subsidy would be smaller relative to the one under the straight salary contract, w. Because workers are risk averse, the risk imposed must be small so that the gains from the subsidy reduction are not outweighed by the cost that risk imposes on risk averse workers.

$$\underline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\underline{\theta}}) \geq \underline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\overline{\theta}}), \tag{$IC\underline{\theta}$}$$

$$\overline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\overline{\theta}}) \geq \overline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\underline{\theta}}), \tag{$IC\overline{\theta}$}$$

$$\theta U(\overline{w}^s_{\theta}) + (1-\theta)U(\underline{w}^s_{\theta}) \geq 0 \text{ for } \theta \in \left\{\overline{\theta}, \underline{\theta}\right\}, \tag{$IR\theta$}$$

$$\theta \overline{w}^s_{\theta} + (1-\theta)\underline{w}^s_{\theta} \leqq y(\theta) \text{ for } \theta \in \left\{\overline{\theta}, \underline{\theta}\right\}. \tag{$ZP\theta$}$$

In the appendix, it is shown formally that the optimal contract $C^s_{\underline{\theta}}$ is such that $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}} = y(\underline{\theta})$, and $C^s_{\overline{\theta}}$ is such that $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ and $\overline{\theta}\overline{w}^s_{\overline{\theta}} + (1-\overline{\theta})\underline{w}^s_{\overline{\theta}} = y(\overline{\theta})$. However, it can be easily seen by adding the incentive compatibility constraints that a necessary condition for a separating equilibrium to exist is that $U(\overline{w}^s_{\overline{\theta}}) - U(\underline{w}^s_{\overline{\theta}}) \geq U(\overline{w}^s_{\underline{\theta}}) - U(\underline{w}^s_{\underline{\theta}})$. That is, the contract tailored to high-ability workers offer at least as much pay-for-performance as the contract tailored to low-ability workers. Hence, if it is optimal to pay a straight salary to a low-ability worker, the contract offered to high-ability workers has to involve pay-for-performance

Similarly, in a pooling equilibrium firms maximize high-ability workers' expected payoff subject to low-ability workers' participation constraint, $(IR\underline{\theta})$, and a zero expected profit constraint when evaluated at the population average probability of success, $\widehat{\theta}$. We will define the contract $C^p \equiv (\overline{w}^p, \underline{w}^p)$ as the solution to the following program,

$$\max_{\{\overline{w}^p, \underline{w}^p\}} \overline{\theta}U(\overline{w}^p) + (1-\overline{\theta})U(\underline{w}^p) \tag{PP}$$

subject to

$$\underline{\theta}U(\overline{w}^p) + (1-\underline{\theta})U(\underline{w}^p) \geq U(0), \tag{$IR\underline{\theta}$}$$

$$\widehat{\theta}\overline{w}^p + (1-\widehat{\theta})\underline{w}^p \leqq \mu y(\overline{\theta}) + (1-\mu)\, y\left(\underline{\theta}\right). \qquad (ZP\widehat{\theta})$$

In the appendix, it is shown formally that the optimal contract C^p is such that $\overline{w}^p > \overline{w}^p$ and $\widehat{\theta}\overline{w}^p + (1-\widehat{\theta})\underline{w}^p = \mu y(\overline{\theta}) + (1-\mu)\, y\left(\underline{\theta}\right)$.

Before stating the main result of this section in a proposition, we will define the population share of high-ability workers, $\widetilde{\mu}$, as the share of high-ability workers for which the solution to PP leaves a high-ability worker at least as well off as she would be under the contract $C^s_{\overline{\theta}}$ and leaves a low-ability worker better off than under $C^s_{\underline{\theta}}$. It takes several steps of simple algebra to show $\widetilde{\mu} = \frac{1}{1+\frac{\overline{\theta}(1-\overline{\theta})}{\triangle\theta}R}$, where $\triangle\theta = \overline{\theta} - \underline{\theta}$ and $R = U'(w^s_{\underline{\theta}})[\frac{1}{U'(\overline{w}^s_{\overline{\theta}})} - \frac{1}{U'(\underline{w}^s_{\overline{\theta}})}] \geq 0$. Given this the following result is formally shown in the appendix.

Proposition 1 *(i) If $\mu \leqq \widetilde{\mu}$, then firms offer a straight-salary contract with $w^s_{\underline{\theta}} = y\left(\underline{\theta}\right)$ and the pay-for-performance contract $\left(\overline{w}^s_{\overline{\theta}}, \underline{w}^s_{\overline{\theta}}\right)$, with $\overline{\theta}\overline{w}^s_{\overline{\theta}} + (1-\overline{\theta})\underline{w}^s_{\overline{\theta}} = y\left(\overline{\theta}\right)$; and (ii) if $\mu > \widetilde{\mu}$, then firms offer the pay-for-performance contract $(\overline{w}^p, \underline{w}^p)$, with $\widehat{\theta}\overline{w}^p + (1-\widehat{\theta})\underline{w}^p = y\left(\widehat{\theta}\right)$ and both types of workers participate.*

The following corollary follows immediately from proposition 1.

Corollary 1 $\overline{w}^s_{\overline{\theta}} - \underline{w}^s_{\overline{\theta}} > \overline{w}^p - \underline{w}^p > \overline{w}^s_{\underline{\theta}} - \underline{w}^s_{\underline{\theta}} = 0.$

This corollary states that the pay-for-performance sensitivity is larger when sorting is achieved. The intuition is straightforward. The pay-for-performance sensitivity needs to be larger in separating equilibrium since low-ability workers must be stopped from mimicking high-ability workers and high-ability workers' payoff in a pooling equilibrium must be larger than in the separating equilibrium. Given that in a pooling equilibrium a high-ability worker's expected compensation is lower than in a separating equilibrium, the only way that his payoff is larger is by mean of a lower pay-for-performance sensitivity.

It is worthwhile to comment on the robustness of this result to more than two types. It is easy to extend the analysis to more than two ability types. If we restrict ourselves to either pure separating or pure pooling equilibrium only, with more than two types the analysis yields the following. In the separating equilibrium only the lowest-ability type receives a straight salary, while all other ability types are paid pay-for-performance. Because of the single-crossing property, the wage spread, $\overline{w}^s_\theta - \underline{w}^s_\theta$, is increasing in workers' ability; that is, for any two workers of ability θ and θ' with $\theta > \theta'$ the wage spread $\overline{w}^s_\theta - \underline{w}^s_\theta$ from the contract tailored to a θ-worker is larger than the wage spread $\overline{w}^s_{\theta'} - \underline{w}^s_{\theta'}$ from the contract tailored to a θ'-worker. This readily follows from the sum of a θ-worker and a θ'-worker's incentive compatibility constraints, which results in that $U(\overline{w}^s_\theta) - U(\underline{w}^s_\theta) \geq U(\overline{w}^s_{\theta'}) - U(\underline{w}^s_{\theta'})$ must be satisfied. In the pooling equilibrium all firms offer the same pay-for-performance contract and workers are indifferent among firms. The pay-for-performance contract offered in equilibrium is the one that the highest-ability workers prefer the most among all those contracts that break even on the population average probability of success, and is accepted by all workers. As in the two type case, a unique pooling equilibrium exists when each worker's ability type prefers the pooling equilibrium to the separating equilibrium, with strict preferences for at least one worker's type. Results, therefore, similar to the ones in proposition 1 hold even with more than two ability types. Whereas when partial pooling is allowed, it is easy to show that partial pooling may be an equilibrium. Suppose, for instance, that there are 3 types, denoted by θ_n, with $\theta_1 < \theta_2 < \theta_3$ and proportions μ_1, μ_2 and μ_3. Then, if μ_2 is sufficiently large and μ_3 is sufficiently small, in equilibrium all workers with an ability equal to or lower than θ_2 are pooled under the pay-for-performance contract that maximizes a θ_2-worker's expected payoff subject to that firms break even, while a θ_3-worker is offered a different pay-for-performance contract satisfying $\overline{w}^s_3 - \underline{w}^s_3 > \overline{w}^p - \underline{w}^p$.

4 Extensions

4.1 Monitoring

One of the main problems of implementing a pay-for-performance contract as the one implemented in the section above is the existence of a good monitoring system able to yield a correct performance measure. In general, however, most monitoring systems are imperfect and measure output with error. In this section, then we analyze the same problem studied above, but assuming that the monitoring technology available allows firms to choose the monitoring precision.

So far we have assumed that monitoring is perfect and costless, while in this section it is assumed that output is correctly observed with probability p, $p \in [\frac{1}{2}, 1]$, and incorrectly with probability $1-p$; *i.e;* p is the monitoring precision. The cost of measuring output with precision p is given by $\gamma C(p)$, where γ captures the monitoring difficulty and the following assumption concerning $C(\cdot)$ is made.

- Assumption: (i) $C\left(\frac{1}{2}\right) \geq 0$ and $C(1) < \infty$; (ii) $C'(p) > 0$ and $C''(p) \geq 0$.

We also assume that firms can commit to a level of monitoring intensity p that is part of the contract. Thus, p is known to the worker at the time to sign a contract. The timing is the same as before, but now a contract specifies not only a wage $\overline{w}$ to be paid when the measured output is high and a wage $\underline{w}$ to be paid when the measured output is low, but also a monitoring intensity p.

Denoting by $\hat{y}$ the observed output, a θ-worker's probability of success, given monitoring precision p, is $p\theta + (1-p)(1-\theta)$ and is denoted by $\theta(p)$. Notice that the first term corresponds to the probability that a θ-worker produces the high output and this is measured correctly, while the second term corresponds to the probability that the low output is produced and this is measured incorrectly; that is, is measured as high-output.[9]

Notice that under this monitoring technology a θ-worker's probability of success is larger than the true probability of success when $\theta > \frac{1}{2}$, while it is smaller when $\theta \leq \frac{1}{2}$. Thus, an increase in the monitoring precision increases a θ-worker's probability of success when $\theta \leq \frac{1}{2}$ and decreases it when $\theta > \frac{1}{2}$. Furthermore, the probability of success for a $\overline{\theta}$-worker increases more with p and is larger than the probability of success for a $\underline{\theta}$-worker for any $p > \frac{1}{2}$. Technically, the monitoring technology assumed guarantees that the single-crossing property holds for any $p > \frac{1}{2}$ and therefore sorting is possible. Finally, when $p = \frac{1}{2}$, which represents no monitoring, the signal is completely uninformative because, regardless of a worker's ability, the probability of success is the same for both, high- and low-ability workers. This implies that while high-ability workers are more likely to produce the high output, both types of workers are equally likely to benefit from a high wage attached to a high-output measure. Therefore, workers cannot be sorted out by mean of a pay-for-performance contract.

Next notice that in any separating equilibrium low-ability workers get at least a payoff equal to $U\left(y\left(\underline{\theta}\right)\right)$. The reason being that when firms offer a contract that yields an expected payoff $\underline{\theta} U\left(\overline{w}\right) + \left(1 - \underline{\theta}\right) U\left(\underline{w}\right)$ lower than $U\left(y\left(\underline{\theta}\right)\right)$, there is a firm that has an incentive to deviate offering a fixed-wage contract that pays $y\left(\underline{\theta}\right) - \varepsilon$. For it attracts all low-ability workers and make positive profits.[10]

Given this, consider the following contract

$$\left(\overline{w}_{\theta}^{sm}, \underline{w}_{\theta}^{sm}, p_{\theta}\right) \equiv \arg\max \overline{\theta}\left(p_{\overline{\theta}}\right) U(\overline{w}_{\overline{\theta}}^{sm}) + (1 - \overline{\theta}\left(p_{\overline{\theta}}\right)) U(\underline{w}_{\overline{\theta}}^{sm}) \qquad (MSP)$$

[9] It also worthwhile to remark that if $p = 1$ and $\gamma = 0$, then the model is as the no-monitoring model. Thus, the no-monitoring model corresponds to the special case of the monitoring model in which monitoring is perfect and costless.

[10] Formally, the proof is exactly as the proof of lemma 4 in the appendix.

subject to

$$\underline{\theta}\left(p_{\underline{\theta}}\right)U(\overline{w}^{sm}_{\underline{\theta}})+(1-\underline{\theta}\left(p_{\underline{\theta}}\right))U(\underline{w}^{sm}_{\underline{\theta}})\geq\underline{\theta}\left(p_{\overline{\theta}}\right)U(\overline{w}^{sm}_{\overline{\theta}})+(1-\underline{\theta}\left(p_{\overline{\theta}}\right))U(\underline{w}^{sm}_{\overline{\theta}}),\qquad(MIC\underline{\theta})$$

$$\overline{\theta}\left(p_{\overline{\theta}}\right)U(\overline{w}^{sm}_{\overline{\theta}})+(1-\overline{\theta}\left(p_{\overline{\theta}}\right))U(\underline{w}^{sm}_{\overline{\theta}})\geq\overline{\theta}\left(p_{\underline{\theta}}\right)U(\overline{w}^{sm}_{\underline{\theta}})+(1-\overline{\theta}\left(p_{\underline{\theta}}\right))U(\underline{w}^{sm}_{\underline{\theta}}),\qquad(MIC\overline{\theta})$$

$$\theta\left(p\right)U\left(\overline{w}^{sm}_{\theta}\right)+(1-\theta\left(p\right))U\left(\underline{w}^{sm}_{\theta}\right)\geq 0\qquad(MIR\theta)$$

$$\theta\left(p\right)\overline{w}^{sm}_{\theta}+(1-\theta\left(p\right))\underline{w}^{sm}_{\theta}\leq y\left(\theta\right)-\gamma C\left(p\right).\qquad(MZP\theta)$$

As in the model with perfect and costless monitoring, it readily follows from the sum of $MIC\underline{\theta}$ and $MIC\overline{\theta}$ that a necessary condition to stop a low-ability worker from mimicking a high-ability worker is that $U(\overline{w}^{sm}_{\overline{\theta}})-U(\underline{w}^{sm}_{\overline{\theta}})\geq U(\overline{w}^{sm}_{\underline{\theta}})-U(\underline{w}^{sm}_{\underline{\theta}})$; that is, the contract tailored to a high-ability worker must involve more risk than the contract tailored to a low-ability worker. This coupled with the fact that competition forces firms to offer low-ability workers an expected payoff at least as large as $U\left(y\left(\underline{\theta}\right)\right)$ implies that in a separating equilibrium two contracts are offered, a straight-salary one that pays $y\left(\underline{\theta}\right)$ and a pay-for-performance contract that satisfies $MIC\underline{\theta}$ and $MZP\overline{\theta}$ with equality. Given that in a separating equilibrium low-ability workers are offered a straight-salary contract, there is no benefit from investing in monitoring; that is, $p_{\underline{\theta}}=\frac{1}{2}$. Also, notice that separation requires that $p_{\overline{\theta}}>\frac{1}{2}$, otherwise any contract that yields high-ability workers a payoff larger than $U\left(y\left(\underline{\theta}\right)\right)$, it also yields low-ability workers a payoff larger than $U\left(y\left(\underline{\theta}\right)\right)$. Thus, $MIC\underline{\theta}$ is violated. The reason being that when $p_{\overline{\theta}}=\frac{1}{2}$, high- and low-ability worker's probability of success is the same.

Consider next the case in which firms' menus have only one contract and define the contract $C^{pm} \equiv (\overline{w}^{pm}, \underline{w}^{pm}, p^p)$ as the solution to

$$(\overline{w}^{pm}, \underline{w}^{pm}, p^p) \equiv \arg\max \overline{\theta}\,(p)\, U(\overline{w}^{pm}) + (1 - \overline{\theta}\,(p))U(\underline{w}^{pm}) \qquad (MPP)$$

subject to

$$\underline{\theta}\,(p)\, U(\overline{w}^{pm}) + (1 - \underline{\theta}\,(p))\, U(\underline{w}^{pm}) \geq 0, \qquad (MIR\underline{\theta})$$

$$\hat{\theta}\,(p)\, \overline{w}^{pm} + (1 - \hat{\theta}\,(p))\underline{w}^{pm} \leq \mu y\left(\overline{\theta}\right) + (1 - \mu)\, y\left(\underline{\theta}\right) - \gamma C\,(p)\,, \qquad (MZPA)$$

where $\hat{\theta}\,(p) = p\left(\mu\overline{\theta} + (1 - \mu)\,\underline{\theta}\right) + (1 - p)\left(1 - \mu\overline{\theta} - (1 - \mu)\,\underline{\theta}\right).$

Notice that as long as $p^p > \frac{1}{2}$, the contract that solves this program is a pay-for-performance contract, while if $p^p = \frac{1}{2}$, the contract that solves this program is a straight-salary contract. The reason being that at $p = \frac{1}{2}$, the average probability of success $\hat{\theta}\,(p)$ is the same as high-ability workers' probability of success. As expected, in the appendix it is shown that if the monitoring difficulty is not too large, $\gamma \leq \gamma^p$, the contract involves pay-for-performance and positive monitoring, while if $\gamma > \gamma^p$, the contract involves no monitoring and a straight-salary. The reason being that the decrease in the subsidy from high-ability workers to low-ability workers that results from a pay-for-performance contract is outweighed by the cost of measuring output with a positive precision.

We define the population share of high-ability workers $\widetilde{\mu}\,(\gamma)$, as the share of high-ability workers for which the solution to MPP leaves a high-ability worker at least as well off as she would be under the contract $C^{sm}_{\overline{\theta}}$ and leaves a low-ability worker better-off than under $C^{sm}_{\underline{\theta}}$.[11] Then, the discussion up to here can be summarized in the next proposition.

[11]It takes several steps of simple algebra to show that $\widetilde{\mu}\,(\gamma) = \frac{1}{1+\frac{\overline{\theta}\left(p_{\overline{\theta}}\right)(1-\overline{\theta}\left(p_{\overline{\theta}}\right))}{\triangle\theta\left(1-2p_{\overline{\theta}}\right)}R\left(p_{\overline{\theta}}\right)}$, where $R\,(\gamma) = U'(w^{sm}_{\underline{\theta}})[\frac{1}{U'(\overline{w}^{sm}_{\overline{\theta}})} - \frac{1}{U'(\underline{w}^{sm}_{\overline{\theta}})}] \geq 0.$

Proposition 2 *(i) If* $\mu \leqq \widetilde{\mu}(\gamma)$*, firms offer a straight-salary contract with* $w^s_{\underline{\theta}} = y(\underline{\theta})$ *and* $p_{\underline{\theta}} = \frac{1}{2}$ *and the pay-for-performance contract* $\left(\overline{w}^s_{\overline{\theta}}, \underline{w}^s_{\overline{\theta}}, p_{\overline{\theta}}\right)$*, with* $\overline{\theta}(p_{\overline{\theta}})\overline{w}^s_{\overline{\theta}} + (1 - \overline{\theta}(p_{\overline{\theta}}))\underline{w}^s_{\overline{\theta}} = y\left(\overline{\theta}\right) - \gamma C(p_{\overline{\theta}})$ *and* $p_{\overline{\theta}} > \frac{1}{2}$*. Low-ability workers choose the straight-salary contract, while high-ability workers choose the pay-for-performance contract; (ii) if* $\mu > \widetilde{\mu}(\gamma)$ *and* $\gamma \leq \gamma^p$*, firms offer the pay-for-performance contract* $(\overline{w}^p, \underline{w}^p, p^p)$*, with* $\widehat{\theta}(p^p)\overline{w}^p + (1 - \widehat{\theta}(p^p))\underline{w}^p = y\left(\widehat{\theta}(p^p)\right) - \gamma C(p^p)$ *and* $p^p > \frac{1}{2}$ *and both types of workers participate; and (iii) if* $\mu > \widetilde{\mu}(\gamma)$ *and* $\gamma > \gamma^p$*, firms offer the straight-salary contract, with* $\overline{w}^p = \underline{w}^p = y\left(\widehat{\theta}\right)$ *and* $p^p = \frac{1}{2}$ *and both types of workers participate*

In short this proposition states that as long as the monitoring difficulty is not too large, results similar to the ones in proposition 1 hold under imperfect and costly monitoring. In fact, it can be shown that proposition 1 corresponds to the special case in which $\gamma = 0$ since $p_{\overline{\theta}}$ and p^p goes to 1 and therefore $\overline{\theta}(1) = \overline{\theta}$ and $\underline{\theta}(1) = \underline{\theta}$. Whereas when the monitoring difficulty is large, a new equilibrium arises, which involves to pay all workers straight salaries. Thus, the model with monitoring predicts that only pay-for-performance contracts should be observed when the monitoring difficulty is small and the proportion of high-ability workers is large; only straight salaries should be observed when the monitoring difficulty and the proportion of high-ability workers are large; and that pay-for-performance and straight salaries should coexist when the proportion of high-ability workers is small.

Furthermore, the model predicts that it is optimal, for certain parameter values, to monitor workers and that to sort workers out is both, pay-for-performance and monitoring should take place. Thus, monitoring and pay-for-performance are complements devices to achieve perfect sorting.

The paper closest to ours is Lazear (1986). Basically, he studies piece-rates and straight salaries in variety of different settings. The section of Lazear's paper that assumes as we do that workers know their own ability and firms do not, shows that under perfect and exogenous, but costly monitoring of output low-ability workers select straight-salary firms, while high-ability workers select piece-rate firms. In Lazear's paper, the piece-rate and straight salary are set to satisfy a zero profit condition and he states that under this assumption his model implies that a straight-salary firm always exists because it will always be able to attract some workers.[12]

We share the sorting result with Lazear, and so our model can be seen as complementary to his. However, we do not need positive monitoring costs for this result and we explicitly model firms as profit maximizers and show that only under certain conditions both compensations methods coexist. Contrary to Lazear's paper, we show that his result that a straight-salary firm always exists is not robust to profit maximizing firms when the proportion of high-ability workers is large and the monitoring difficulty is not too large because in that case only pay-for-performance firms are observed. Thus, we show that the existence of positive monitoring costs is neither necessary nor sufficient for the co-existence of pay-for-performance and straight-salary contracts; that is, even when there is no monitoring cost, there is a separating equilibrium for certain parameter values and when there are positive monitoring costs, there might be an equilibrium in which only pay-for-performance firms exist. Furthermore, we show that competing firms in equilibrium choose to monitor workers, but imperfectly unless the monitoring difficulty is small. The reason for investing in monitoring is to decrease the intensity of competition and offer high-ability workers a mix between pay-for-performance sensitivity and monitoring intensity that either decreases a high-ability worker's cost of being separated from low-ability workers or decreases the subsidy that high-ability workers give to low-ability workers when they are pooled together. Hence, our model is similar in spirit to Lazear's since we both use a sorting model to explain why some firms offer pay-for-performance and others straight salaries, but contrary

[12] This result is due to positive monitoring costs. If these were zero, then all firms offer a pay-for-performance; that is, each worker gets pay his realized output.

to Lazear's model in our model there are cases in which no straight-salary firm exists and cases in which only straight-salary exists. In addition, we show that, unless the monitoring difficulty is very small, firms never choose to monitor output perfectly. These are important difference because the empirical evidence as casual observations shows that for many occupations, like sale workers and CEOs, only pay-for-performance workers are observed, while in others, like agricultural workers, straight-salary and pay-for-performance workers are observed and that firms in which the monitoring difficulty is large are much less likely to monitor and pay-for-performance (MacLeod and Parent, 1998; Brown, 1990).

In addition, the difference in the modeling assumptions allows us to make predictions concerning the optimal pay-for-performance sensitivity, which is always equal to 1 in Lazear's model, the monitoring intensity and help us to identify the conditions under which different types of compensation methods should be observed.

4.2 Effort

In general, it is argued that pay-for-performance contracts are designed to induce additional employee effort, increase production and, as a result, compensation. While it may seem obvious that a pay-for-performance contract induces more effort from workers, it is not. When a firm chooses a straight salary compensation regime, usually conditions the wage to a minimum level of output. Hence, it would be possible that the minimum required in a straight-salary job is larger than the minimum required in a pay-for-performance job. Furthermore, it may be the case that the minimum required output can be accomplished only by the most able workers and when pay-for-performance and straight salary coexist, more heterogeneous minimum can be accepted, resulting in lower levels of output.

So far, we have assumed that output is independent of effort, while in this section we assume that output can take on two values $\bar{y}(e)$ and $\underline{y}(e)$, with $\bar{y}(e) - \underline{y}(e) > 0$ and $y_e(e) > 0$

for all $e \in \Re_+$, where e is an unobservable effort $y_e(\cdot)$ denotes the derivative of output with respect to effort.[13] A θ-worker's cost of exerting e units of effort is independent of her ability and equal to e. As before a θ-worker's probability of producing the high output for any given effort is θ. Hence, a θ-worker's expected output when she exerts an effort level equal to e, denoted by $y(e,\theta)$, is $\theta\bar{y}(e) + (1-\theta)\underline{y}(e)$.

- Assumption 4: $y_e(e) > 0$, $y_{ee}(e) \leqq 0$ and $\overline{y}_e(e) \geq \underline{y}_e(e)$ for all $e \in [0, E]$.

Each firm offers a menu of contracts, where each contract C_k specifies a non-negative wage paid for each output level; *i.e.*, $\overline{w}_k$ for $\bar{y}_k$ and $\underline{w}_k$ for $\underline{y}_k$, output levels $\bar{y}_k$ and $\underline{y}_k$ and a payment B_k for $y_k \notin \left\{\underline{y}_k, \bar{y}_k\right\}$; that is, contract C_k is given by $\left(\overline{w}_k, \underline{w}_k, \bar{y}_k, \underline{y}_k, B_k\right)$. Furthermore, we assume that U is unbounded from below; that is, $U(w) \longrightarrow -\infty$ as $w \longrightarrow \hat{w}$, where $\hat{w} \geq -\infty$ and the timing is as before, but now after the contract has been signed, workers exert effort, output is realized and compensation takes place.

It is easy to show that a θ-worker's optimal effort level under full-information, denoted by e^*_θ is determined by the following first order condition

$$U'(y(e^*_\theta, \theta))\, y_e(e^*_\theta, \theta) - 1 = 0. \tag{3}$$

It follows from 3 and assumption 4 that $y(e^*_{\overline{\theta}}, \overline{\theta}) \geq y(e^*_{\underline{\theta}}, \underline{\theta})$. Therefore, under full information $e^*_{\overline{\theta}} \lesseqqgtr e^*_{\underline{\theta}}$. Furthermore, it is assumed that $U(y(e^*_\theta, \theta)) - e^*_\theta > 0$ for $\theta \in \left\{\overline{\theta}, \underline{\theta}\right\}$; that is, efficient for both ability types to participate and exert the optimal effort level.

Notice that the market can not implement the first-best level of efforts $(e^*_{\overline{\theta}}, e^*_{\underline{\theta}})$ under a straight-salary regime and forcing contracts.[14] To see this suppose that firms offer the following contracts $C^*_{\underline{\theta}} = \left(y(e^*_{\underline{\theta}}, \underline{\theta}), y(e^*_{\underline{\theta}}, \underline{\theta}), \bar{y}^*_{\underline{\theta}}, \underline{y}^*_{\underline{\theta}}, B_{\underline{\theta}}\right)$ and $C^*_{\overline{\theta}} = \left(y(e^*_{\overline{\theta}}, \bar{\theta}), y(e^*_{\overline{\theta}}, \bar{\theta}), \bar{y}^*_{\overline{\theta}}, \underline{y}^*_{\overline{\theta}}, B_{\overline{\theta}}\right)$.

[13] The results of this section do not depend on the assumption that output is a binary random variable. They hold for any density function that has a lower bound and movable support because the firm can always detect a deviation outside of the set of required outputs and punish such deviation with a large penalty.

[14] By forcing contracts we mean that when $y_\theta \notin \left\{\underline{y}_\theta, \bar{y}_\theta\right\}$ the agent can be severely punished, so that the principal can restrict attention to outputs in $\left\{\underline{y}_\theta, \bar{y}_\theta\right\}$.

Because $y(e,\bar{\theta}) > y(e,\underline{\theta})$ for all e and $e^*_\theta = \arg\max\{U(y(e_\theta,\theta)) - e_\theta\}$, $U(y(e^*_{\overline{\theta}},\bar{\theta})) - e^*_{\overline{\theta}} > U(y(e^*_{\underline{\theta}},\bar{\theta})) - e^*_{\underline{\theta}} > U(y(e^*_{\underline{\theta}},\underline{\theta})) - e^*_{\underline{\theta}}$. Therefore, a $\underline{\theta}$-worker has incentive to claim to be a $\overline{\theta}$-worker by choosing $e^*_{\overline{\theta}}$ and taking the contract designed for a high-ability worker. Thus, firms are faced with a similar problem to the one in section 3.

As in section 3, it can be shown that there are two types of equilibrium, a separating and a pooling equilibrium. The proof consists of minor extensions of the one for the no-effort model, therefore for the sake of brevity it will be omitted.[15] Nevertheless, in the next proposition we characterize the optimal contract for the case in which the equilibrium is separating and for the case in which is pooling. Before doing so, it is useful to define the function $h(w) \equiv \frac{1}{U'(w)}$ and the following contracts:

$$(\overline{w}^{se}_\theta, \underline{w}^{se}_\theta, e^s_\theta)^{\overline{\theta}}_{\underline{\theta}} \equiv \arg\max \overline{\theta}U(\overline{w}^{se}_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^{se}_{\overline{\theta}}) - e^s_{\overline{\theta}} \qquad (EPS)$$

subject to

$$\underline{\theta}U(\overline{w}^{se}_{\underline{\theta}}) + (1-\underline{\theta})U(\underline{w}^{se}_{\underline{\theta}}) - e^s_{\underline{\theta}} \geq \underline{\theta}U(\overline{w}^{se}_{\overline{\theta}}) + (1-\underline{\theta})U(\underline{w}^{se}_{\overline{\theta}}) - e^s_{\overline{\theta}}, \qquad EIC\underline{\theta}$$

$$\overline{\theta}U(\overline{w}^{se}_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^{se}_{\overline{\theta}}) - e^s_{\overline{\theta}} \geq \overline{\theta}U(\overline{w}^{se}_{\underline{\theta}}) + (1-\overline{\theta})U(\underline{w}^{se}_{\underline{\theta}}) - e^s_{\underline{\theta}}, \qquad (EIC\overline{\theta})$$

$$\underline{\theta}U(\overline{w}^{se}_{\underline{\theta}}) + (1-\underline{\theta})U(\underline{w}^{se}_{\underline{\theta}}) - e^s_{\underline{\theta}} \geq 0, \qquad (EIR\underline{\theta})$$

$$\theta\overline{w}^{se}_\theta + (1-\theta)\underline{w}^{se}_\theta \leqq y(e^s_\theta,\theta), \text{ for } \theta \in \left\{\underline{\theta},\overline{\theta}\right\}. \qquad (EZP\theta)$$

and

$$(\overline{w}^p, \underline{w}^p, e^p) \equiv \arg\max \overline{\theta}U(\overline{w}^p) + (1-\overline{\theta})U(\underline{w}^p) - e^p \qquad (EPP)$$

[15] The proof is available upon request.

subject to

$$\underline{\theta}U(\overline{w}^p)+(1-\underline{\theta})U(\underline{w}^p)-e^p\geq 0, \qquad EIR\underline{\theta}$$

$$\widehat{\theta}\overline{w}^p+(1-\widehat{\theta})\underline{w}^p\leqq y\left(e^p,\widehat{\theta}\right). \qquad (EZP\widehat{\theta})$$

Notice that we have ignored the incentive compatibility constraints for effort. The reason being that $U(B_\theta)\longrightarrow-\infty$ as $B_\theta\longrightarrow\hat{B}_\theta$, where $\hat{B}_\theta\geq-\infty$. Thus, it is always possible to choose a penalty B_θ sufficiently severe that a θ-worker has no incentive to produce an output $y_\theta\notin\left\{\underline{y}_\theta,\bar{y}_\theta\right\}$. This leads to the following result.

Proposition 3 *(i) If $\mu\leqq\widetilde{\mu}$, then firms offer the straight-salary contract $C^s_{\underline{\theta}}\left(e^*_{\underline{\theta}}\right)$ and the pay-for-performance contract $C^s_{\overline{\theta}}\left(e^s_{\overline{\theta}}\right)$, with effort level $e^s_{\overline{\theta}}<e^*_{\overline{\theta}}$ if $h(\cdot)$ is convex and $e^s_{\overline{\theta}}\geq e^*_{\overline{\theta}}$ if $h(\cdot)$ is concave.*[16] *Low-ability workers choose the straight-salary contract $C^s_{\underline{\theta}}\left(e^*_{\underline{\theta}}\right)$, while high-ability workers choose the pay-for-performance contract $C^s_{\overline{\theta}}\left(e^s_{\overline{\theta}}\right)\equiv\left(\overline{w}^{se}_{\overline{\theta}},\underline{w}^{se}_{\overline{\theta}},e^s_{\overline{\theta}}\right)$; (ii) If $\mu>\widetilde{\mu}$, then every firm offers the pay-for-performance contract $C^p(e^p)\equiv(\overline{w}^p,\underline{w}^p,e^p)$, with $e^p\leq e^*_{\underline{\theta}}$ and $e^*_{\overline{\theta}}\gtreqqless e^p$ if $h(\cdot)$ is concave, and $e^p\geq e^*_{\overline{\theta}}$ and $e^*_{\underline{\theta}}\gtreqqless e^p$ if $h(\cdot)$ is convex, and both types of workers participate.*

This proposition states that the results in proposition 1 are robust to the introduction of effort when workers have unlimited liability. That is, when the equilibrium is separating, low-ability workers choose a straight-salary job and exert the efficient level of effort, while high-ability workers choose a pay-for-performance job and exert, in general an inefficient level of effort. Whereas, when the equilibrium is pooling both, high- and low-ability workers work in a pay-for-performance job and exert an inefficient level of effort.

[16] $h''\gtreqqless 0$ if and only if $U'''\lesseqqgtr 2\frac{U''^2}{U'}$.

The intuition for a separating equilibrium follows from that low-ability workers can be discouraged from mimicking high-ability workers by either imposing a large amount of risk and a low performance standard or by a small amount of risk and a high performance standard. For a given performance standard; *i.e.,* level of effort, when workers are highly risk averse (h is convex) a small increase in risk is needed to stop low-ability workers from mimicking high-ability workers, while when workers are not as risk averse (h is concave) a large increase in risk is needed to induce workers self-selection. Therefore, in order to stop a low-ability worker from mimicking a high-ability worker at the minimum cost possible for high-ability workers, firms offer a contract that sets a high (low) performance standard and offers an small (larger) amount of risk when workers are highly (not as highly) risk averse; that is, firms trade-off performance standards and risk to minimize high-ability workers' lost in well-being due to the existence of asymmetric information. Thus, the model predicts that in separating equilibrium a trade-off between pay-for-performance sensitivity and effort different from the one predicted by the moral hazard model exits; that is, workers under contracts with larger pay-for-performance sensitivity exert less effort, while workers with smaller pay-for-performance sensitivity exert more effort.

The intuition for the existence of an equilibrium in which all firms offer the same pay-for-performance contract when the proportion of high-ability workers is large is the same as the one in the pure asymmetric information model with no effort. That is, when the combined output of all agents is such that firms can offer a contract that make both, high- and low-ability workers better-off than they would be, were firms offer a menu of contracts that separate types, it is optimal to pool all types under the same contract.

There are two interesting corollaries that readily follows from proposition 3. One concerning the standard argument that pay-for-performance contracts are designed to induce additional employee effort, increase production and, as a result, compensation, and another one, concerning the optimal levels of effort in a separating equilibrium.

Corollary 2 *Suppose that* $\overline{y}_e(e) = \underline{y}_e(e)$ *for all* e. *Then, (i)* $e^*_{\overline{\theta}} < e^*_{\underline{\theta}}$ *and* $y\left(e^*_{\overline{\theta}}, \overline{\theta}\right) = y\left(e^*_{\underline{\theta}}, \underline{\theta}\right)$*; and (ii) if* h *is convex,* $e^s_{\overline{\theta}} < e^*_{\overline{\theta}}$ *and* $y\left(e^s_{\overline{\theta}}, \overline{\theta}\right) < y\left(e^*_{\underline{\theta}}, \underline{\theta}\right) = y\left(e^s_{\underline{\theta}}, \underline{\theta}\right)$.

This corollary shows that low-ability workers under a straight-salary contract may have a higher performance standard and a larger expected productivity than high-ability workers under a pay-for-performance contract. Thus, in this case pay-for-performance contracts are not designed to induce additional employee effort, increase production and, as a result, compensation. Pay-for-performance contracts are designed together with performance standards to minimizes high-ability workers' cost of separation. Thus, ignoring the sorting effect of pay-for-performance may lead to design incorrect empirical strategies to test the effect of pay-for-performance contracts on productivity.

The next corollary concerns the optimal level of effort in a separating equilibrium.

Corollary 3 *Suppose that* $h''(w) = 0$*, then* $e^s_{\overline{\theta}} = e^*_{\overline{\theta}}$ *and* $e^s_{\underline{\theta}} = e^*_{\underline{\theta}}$.[17]

This proposition establishes that when workers have unlimited liability and the utility function satisfies certain properties, the output standard is set to the first-best level for both, high- and low-ability workers. This shows that pay-for-performance contracts may arise in a world in which a performance standard contract is sufficient to induce the first-best effort level, and pay-for-performance contracts have no effect on effort.[18] This provides a foundation for the simpler model in the main section.

Finally, there are two remarks to the results derived here. First, that the results here are robust to the introduction of more than two types. Second, we have derived our solution assuming that workers have unlimited liability. That is, we have assumed that firms can severely punish a worker whose output is different from the standard set by the firm so that it is never optimal to deviate from that standard.

[17] This is true, for instance, for $U(w) = \alpha \ln(c + w)$, with $c > 0$ and $\alpha > 0$.

[18] This result is similar in spirit to Mirrless (1974), in the sense that the optimal effort approaches the first-best sufficiently close.

The question then becomes how restrictive this assumption is for our solution. To answer this, let suppose for simplicity that a worker cannot be paid in either state a wage such that the utility is lower than 0; that is the maximum punishment that a firm can apply to a worker who claims to be a θ-worker and either overperform or underperform with respect to the performance standard is $\hat{B}_\theta$, where $U\left(\hat{B}_\theta\right) = 0$.[19] Then, it is easy to show that paying low-ability workers a straight salary and demanding their first-best level of effort is incentive compatible if the following holds

$$U\left(y(e^*_{\underline{\theta}}, \underline{\theta})\right) \geq \max\left\{\frac{e^*_{\underline{\theta}} - e^L_{\underline{\theta}}}{(1-\underline{\theta})}, e^*_{\underline{\theta}}\right\}, \tag{4}$$

where e^L_θ solves $\underline{y}(e^L_\theta) = \overline{y}(e^s_\theta)$.

Furthermore, the full solution under the assumption of unlimited liability is implementable under limited liability if together with 4, the following holds[20]

$$\frac{\overline{\theta}U\left(\overline{w}^s_{\overline{\theta}}\right) - \left(e^s_{\overline{\theta}} - e^H_{\overline{\theta}}\left(e^s_{\overline{\theta}}\right)\right)}{(1-\overline{\theta})} \geq \frac{U\left(y(e^*_{\underline{\theta}}, \underline{\theta})\right) - U\left(\underline{w}^s_{\overline{\theta}}\right) + e^s_{\overline{\theta}} - e^*_{\underline{\theta}}}{\underline{\theta}} \geq \max\left\{\frac{\left(e^s_{\overline{\theta}} - e^L_{\overline{\theta}}\left(e^s_{\overline{\theta}}\right)\right) - (1-\overline{\theta})U\left(\underline{w}^s_{\overline{\theta}}\right)}{\overline{\theta}}, \frac{e^s_{\overline{\theta}} - U\left(\underline{w}^s_{\overline{\theta}}\right)}{\overline{\theta}}\right\}, \tag{5}$$

where e^H_θ solves $\underline{y}(e^H_\theta) = \overline{y}(e^s_\theta)$.[21]

Therefore, if 4 and 5 hold for $U\left(\hat{B}_{\overline{\theta}}\right) = 0$, limited liability does not necessarily prevent forcing contracts under unlimited liability from being optimal. Moreover, if $U(0) = 0$, this implies that the threat of penalty is always credible, for the firm can simply refuse to pay

[19] The analysis here also applies for a restriction on the minimum utility different from 0.

[20] Here we are assuming that $\max\left\{(1-\overline{\theta})U(\overline{w}^s_{\overline{\theta}}) - e^H_{\overline{\theta}}, \overline{\theta}U(\underline{w}^s_{\overline{\theta}}) - e^L_{\overline{\theta}}\right\} \geq 0$.

[21] In a pooling equilibrium, the solution under unlimited liability is optimal under limited liability when

$$\theta U(\overline{w}^p) + (1-\theta)\, U(\underline{w}^p) - e^p \geq \max\left\{(1-\theta)U(\overline{w}^p) - e^H\left(e^p\right), \theta U(\underline{w}^p) - e^L\left(e^p\right), 0\right\}, \text{ for } \theta \in \left\{\overline{\theta}, \underline{\theta}\right\}, \tag{6}$$

where e^H solves $f(e^H, \underline{s}) = f(e^p, \overline{s})$ and e^L solves $f(e^L, \overline{s}) = f(e^p, \underline{s})$.

when the output fails to meet the required performance standard. This shows that a bound on the agent's utility does not necessarily preclude the existence of a separating equilibrium in which low-ability workers are paid a straight salary and exert the first-best level of effort. It also follows from 5 that if the first inequality is violated under the optimal contract under unlimited liability, the obedience constraint imposes a limit on how large the pay-for-performance sensitivity can be. This implies that in order to satisfy low-ability workers' incentive compatibility constraint ($IC\underline{\theta}$), the firm must increase the amount of effort required in the contract tailored to high-ability workers. On the other hand, if the second inequality is violated, the obedience constraint imposes a limit on how small the pay-for-performance can be. This implies that in order to satisfy the $IC\underline{\theta}$, the firm must decrease the amount of effort required in the contract tailored to high-ability workers. Hence, the minimum-effort requirement and pay-for-performance sensitivity are substitutes when the limited liability constraint becomes binding.

Therefore, as long as 4 holds, the existence of an effort choice and asymmetric information does not preclude the co-existence of pay-for-performance and straight salary contracts within an occupation. Furthermore, if 5 holds and $h'' = 0$, the existence of asymmetric information does not preclude the implementation of the efficient level of effort. This implies that the loss from these two different problems of information is only the inefficient allocation of risk.

Finally, it is interesting to notice that contrary to the standard literature on screening in a monopolistic setting and with risk-neutral agents, here, the distortion occurs at the top; that is, the higher type's effort is distorted either downwards or upwards. The reason being that with risk-averse agents, there is another instrument to achieve sorting which is the pay-for-performance sensitivity and competition does not allow to distort low-ability workers' performance standards (see, Laffont and Tirole, 1992).

5 Discussion

The basic model of this paper has concrete empirical predictions concerning the type of workers and compensation contracts that should be observed.

The obvious implications of the model is that for an occupation in which the proportion of high-ability workers is small, workers under a straight salary have a lower average quality than workers under pay-for-performance. The best workers, when the equilibrium is separating, select contracts where performance has a payoff and firms aware of that choose contracts accordingly. This implies that productivity is larger among pay-for-performance workers, yet this does not imply that changing all workers to pay-for-performance results in a larger average productivity. In fact, it would have no effect at all on average output, which is the classic screening result.[22] Hence, the productivity gain when both compensation methods coexist is due only to the sorting effect. The evidence is consistent with this prediction. When pay-for-performance and straight-salary contracts coexist within an occupation, pay-for-performance workers earn more and have a higher average productivity than straight salary workers.[23] The compensation and productivity differences ranged roughly from 5 % to 37 % (see, for instance Brown, 1992; Lazear, 1997; Paarsh and Shearer, 1997; Parent and MacLeod, 1998; Petersen, 1991 and 1992; Seiler, 1984). Foster and Rosenzweig (1996) using detailed data from an agricultural labor market in which workers can work in either a piece-rate or a straight-salary occupation find evidence in favor of a one-factor productivity model and that information asymmetries are present,[24] but workers are sorted out according

[22] This may not be true if the participation decision is changed by the change in the compensation method used or if effort is involved.

[23] The standard agency theory cannot fully account for this finding since it cannot explain the coexistence of straight-salary and pay-for-performance contracts within an occupation, but it can explain difference across occupations.

[24] One-factor productivity model means that skills can be summarized in one input (ability) and that the more skilfull workers are more productive accross all occupations.

to their comparative advantages; that is, the more skillful workers work in the piece-rate sector, while the less skillful ones work in the straight-salary sector.

The model, however, predicts a different outcome when the proportion of high-ability workers is large. In this case, all workers are paid pay-for-performance. This is suggestive of when we should observe pay-for-performance and straight salaries within an occupation and when we should observe only pay-for-performance. Clearly, in those occupations in which the pool of applicants has a large share of high-ability workers, we should observe only pay-for-performance workers, while in those in which the share is small, both compensation methods should be observed. As an example of this, in low observable skill occupations like agricultural jobs is common to observe that straight salaries and pay-for-performance contracts coexist, while in occupations that demand higher observable skills, like managerial jobs, only pay-for-performance should be observed. So, as long as observable skills are correlated to unobservable skills, our model suggests an explanation for this phenomenon.

The model not only suggests when different outcomes should be observed, but also makes predictions concerning the magnitude of the pay-for-performance sensitivity in each equilibrium. The model predicts that the pay-for-performance sensitivity in a separating equilibrium is larger than the sensitivity in a pooling equilibrium. There is evidence that provides some support for this prediction. For instance, the pay-for-performance sensitivity for CEOs, occupation in which all workers are paid pay-for-performance, is rather low. Jensen and Murphy (1990) find that CEO's wealth changes $3.25 for every $1000 change in shareholder value, which implies a pay-for-performance sensitivity of *0.003* and conclude that the "the lack of strong pay-for-performance incentives for CEOs indicated by our evidence is puzzling."[25] Whereas in occupations in which both compensation methods coexist, the

[25] Similar evidence is found in Kaplan (1994), Gibbons and Murphy (1990) and Murphy (1985, 1986). Murphy (1999), in a review of the CEO literature, concludes that evidence from several studies and samples leaves us fairly secure that the estimated pay-for-performance sensitivity for CEO's is rather small, between

pay-for-performance sensitivities are much larger. In Lazear's windshield-installing firm the pay-for-performance sensitivity is as high as *0.5* and according to the BLS (1975), mechanics receive usually between 45% to 50% of the labor costs charged to customers and a 50% among agricultural workers is rather common. This is also common in sales-like jobs as in franchises where we can find pay-for-performance sensitivities of 50 percent or more. For instance, in McDonald's case, franchisees pay 5-10 percent royalties on sales, implying an effective commission rates of more than 90 percent, while the company owned stores managers are paid straight salaries.

A shortcoming of the simple model is that in no equilibrium only straight-salary contracts are offered. When monitoring is introduced, the model predicts that only pay-for-performance contracts should be observed when the monitoring difficulty is small and the proportion of high-ability workers is large; only straight salaries should be observed when the monitoring difficulty and the proportion of high-ability workers are large; and that pay-for-performance and straight salaries should coexist when the proportion of high-ability workers is small. Brown (1990), using the industry wage survey, finds that pay-for-performance is positively correlated to the ease of monitoring. MacLeod and Parent (1998) using the PSID, NLSY, CPS and the Quality of Employment Survey, conclude after a very careful analysis that pay-for-performance is more likely to be observed the easiest is to assess workers' performance.[26] Furthermore, as predicted by the model, jobs that pay more, monitors more and jobs that pay less, monitors less (see, for instance, Capelli and Chauvin (1991), Groshen and Kruger (1990) and Neal (1992)).

When effort is considered the model predicts that under certain conditions pay-for-performance workers may not only exert less effort than straight-salary workers, but also have a lower expected productivity. Furthermore, the model predicts that in separating equi-

0.001 and 0.007.

[26] See also Heywood et. al. (1997) and Eisenhardt (1988) and Lafontaine and Slade (1998).

librium a trade-off between pay-for-performance sensitivity and effort different from the one predicted by the moral hazard model exits; that is, workers under contracts with larger pay-for-performance sensitivity exert less effort, while workers with smaller pay-for-performance sensitivity exert more effort. Thus, performance standards and the pay-for-performance sensitivity are substitutes instruments to achieve perfect sorting and provide incentives for effort. Thus, ignoring the sorting effect of pay-for-performance may lead to mistakenly interpret the data on compensation. There is no evidence that we are aware-off that can provide support for this prediction, but it is useful when studying compensation data to have in mind that not always pay-for-performance workers earn and produce more than straight-salary workers.

6 Conclusions

We have developed a model that provides a rationale for the existence of pay-for-performance contracts in the absence of incentive for effort and explains when and in which occupations pay-for-performance is more likely to be observed. In our model pay-for-performance contracts arise for two completely different reasons. When the proportion of high-ability workers is small pay-for-performance contracts arise as sorting device to elicit information about workers' skills, while when the proportion is large pay-for-performance contracts arise to decrease good workers' costs of being pooled with bad workers. Competition among firms for the best workers forces firm to link pay to performance in order to provide the best workers with a higher expected compensation. Furthermore, the model, when effort is considered, predicts, contrary to the moral hazard model, that there is an equilibrium in which workers under contracts with a larger pay-for-performance sensitivity exert less effort than workers under contracts with a smaller pay-for-performance sensitivity. Thus, study incentives and compensation based only on the standard incentive effect paradigm and neglecting the role of competition and asymmetric information, results in a lack of understanding of the role played by incentive contracts and bias the empirical work that tries to identify the effects of contract on productivity and compensation. In short, as the evidence suggests, our paper shows that sorting matters.[27]

The paper also makes contributions to the theoretical literature on screening games. It is shown that in a competitive market and under a slightly modified timing than the one proposed by Rothschild and Stiglitz' (1976) a unique equilibrium exits when a appropiately chosen equilibrium refinement is used. We have shown that when the proportion of high-ability types is sufficiently favorable, the unique equilibrium entails pooling, while under the Rothschild and Stiglitz' timing under the same parametrization no equilibrium exists. In a sense, we have shown that the market is not as good as mechanism reducing the negative externality that low-ability types impose on high-ability types. On the other hand, we have also shown that contrary to the literature on screening in monopolistic settings, under competition the distortion, if any, takes place at the top (see, Laffont and Tirole, 1996 for the standard no distortion at the top result). The reason being that competition forces firms to offer low-ability types the contract that maximizes their expected payoff and therefore, high-ability worker's contract is distorted to stop low-ability workers from mimicking high-ability workers.

While, we have concentrated on compensation methods, the results derived here apply to a wide variety of issues. Consider, for instance, the market for loans. Bester (1985) have shown that a collateral arises as a way to avoid credit rationing and to achieve perfect sorting. We have shown that our model predictions are consistent with this prediction only when the proportion of high-risk borrowers is high. When the proportion of high-risk borrowers is

[27] There is plenty of evidence of the importance of selection due to contract offers. For instance, Lazear (1997) and Paarsh and Shearer (1997) show that almost half of the increase in productivity due to the use of pay-for-performance occurs from a selection effect, and Shearer (1997) shows that the tenure-productivity profile of piece-rate workers is much flatter after controlling for workers' unobserved heterogeneity.

low, a collateral is used to decreases low-risk borrowers' cost of being pooled with high-risk borrowers. So, the presence of a collateral cannot be taken as an evidence that borrowers' types are being sort out. Furthermore, consider also the insurance market. It readily follows from the existence of a pooling equilibrium that pooling risks is be welfare enhancing when the proportion of low-risk agents is large. Also the model predicts that in general it is not optimal for insurance firms to monitor perfectly whether an accident occurred or not. This may explain why in the car insurance industry some accident as an stolen car radio usually do not require that the owner of the car proves that the accident has taken place.

References

[1] Baker, G., Gibbs, M. and Hölmstrom, B., "The Wage Policy of a Firm." *Quarterly Journal of Economics*, February 1994a, vol 109, 922-955.

[2] Baker, G., Gibbs, M. and Hölmstrom, B., "The Internal Economics of the Firm: Evidence from Personnel Data." *Quarterly Journal of Economics*, February 1994b, vol. 109, 881-919.

[3] Balmaceda, F., "Dynamic Job Assignments and Compensation: Theory and Evidence" Mimeo, Cornell University, August, 1997.

[4] Balmaceda, F., "Adverse Selection and Moral Hazard in Competitive Markets" Mimeo, Dartmouth College, October, 1998.

[5] Bester, H., "Screening vs. Rationing in Credit Markets with Imperfect Information" *American Economic Review*, September 1985, vol. 71, 850-55.

[6] Biglaiser, G. and Mezzetti, C., "Principals Competing for an Agent in the Presence of Adverse Selection and Moral Hazrd", *Journal of Economic Theory*, 1993, vol. 61, no. 2: 302-30

[7] Booth, A. and Frank, J., "Productivity, Compensaion and Performance Related Pay" *Journal of Labor Economics* (forthcoming).

[8] Brown, C., "Firms Choice of Method of Pay" *Industrial and Labor Relations Review*, January 1990, vol. 43, no. 1, 165S-182S.

[9] Brown, C., "Wage Levels and the Method of Pay" *Rand Journal of Economics*, Autumn 1992, vol. 23, no. 3, 367-75.

[10] Chiappori. P.A., and Salanie, B., "Empirical Contract Theory: The case of insurance data" *European Economic Review*, 1997, vol. 41, no. 3, 943–50.

[11] Cho, I. and Kreps, D., "Signaling Games and Stable Equilibria" *Quarterly Journal of Economics*, February 1987, vol. 102, 179-222.

[12] Drago, R. and Heywood, J.S., "The Choice of Payment Systems: Australian Establishment Data" *Industrial Relations,* October, 1*995,* Vol. 34, No 4.

[13] Fernie, S. and Metcalf, D., "It's Not What You Pay it's the Way that You Pay It and That's What Gets Results: Jockey's Pay and Performance", *Centre for Economic Performance*, Discussion Paper No. 295, May 1996.

[14] Fernie, S. and Metcalf, D., "(Not) Hanging on the Telephone: Payment Systems in the New Sweatshops", *Centre for Economic Performance*, Discussion Paper No. 310, May 1997.

[15] Ferral, C., "An Empirical Analysis of Occupational Hierarchy" *Journal of human Resources*, 1997, vol. 32, no. 1: 1-34.

[16] Foster, A.D. and Rosenweig, M.R., "Comparative Advantage, Information and the Allocation of Workers to Tasks: Evidence from an Agricultural Labor Market" *Review of Economic Studies*, 1996, vol.63: 347-374.

[17] Foster, A.D. and Rosenweig, M.R., "Information, Learning, and Wage Rates in Low-Income Rural Areas" *Journal of human Resources*, 1993, vol. 28, no. 4: 769-90.

[18] Gibbs, M. "Incentive Compensation in a Corporate Hierarchy." *Journal of Accounting and Economics*, 1995, vol. 19, no. 4: 247-277.

[19] Gibbs, M., and Hendricks, W., "Are Administrative Pay Systems a Veil?" *University of Michigan Business School* Working paper, June 1996.

[20] Gibbons, R.S., and Murphy, K.J., "Relative Performance Evaluation for Chief Executive Office", *Industrial and Labor Relations Review*, January 1990, vol. 43, no. 1, 110-122S.

[21] Hellwig, M., "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection" *European Economic Review*, 1987, vol. 31, no. 3: 319-325.

[22] Heywood, J.S., Siebert, W.S., and Xiandong, W., "Payment by Result Systems: ABritish Evidence" *British Journal of Industrial Relations,* MArch, 1*997,* Vol. 35, No 1: 1-22.

[23] Jensen,M.C., and Murphy, K.J., "Performance Pay and Top Management Incentives", *Journal of Political Economy*, 1990, vol. 98, no. 2: 225-265.

[24] Khalil, M. and Lawaree, M.,"Input versus Outpu Monitoring: Who is the Residual Claimant", *Journal of Economic Theory*, 1995, vol. 78, no. 2: 1-14.

[25] Kaplan, S., "Top Executive Rewards and Firm Performance: A Comparison of Japan and United States", *Journal of Political Economy*, 1994, vol. 102, no. 3: 510-546.

[26] Laffont, J., and J. Tirole, "Adverse Selection and Renegotiation In Procurement." *Review of Economic Studies*, 1990, vol.57: 597-625.

[27] Laffont, J., and J. Tirole, "Adverse Selection and Renegotiation In Procurement." *Econometrica*, 1988, vol.56: 1153-1175.

[28] Lafontaine, F., "Agency Theory and Franchising: Some Emprical Results" *Rand Journal of Economics*, Autumn 1992, vol. 23, no. 3, 263-83.

[29] Lazear, E.P.,"Salaries and Piece-rates", *Journal of Business*, 59: 405-431.

[30] Lazear, E.P.,"Performance and Prouctivity.", NBER Working paper 5672, July 1996.

[31] Lazear, E.P.,"Personnel Economics", Boston, MIT Press, 1995.

[32] Lazear, E.P.,"The Job as a Concept", in *Performance Measurement, Evaluations, and Incentives* Bruns, W. (eds.) Boston, Harvard Business School Press.

[33] MacLeod, B. and Parent, D., "JOb Characteristics and the Form of Compensation" *Cirano* Working Paper 98s-08*,* March, 1*998.*

[34] Mailath, J.M, "Signaling Games" *Recent Developments in Game Theory*, 1993, vol. 60: 241-276.

[35] Mailath, J.M, Okuno-Fujiwarwa, M., and Postlewaite, A., "Belief-Based Refinements in Signaling Games" *Journal of Economic Theory*, 1993, vol. 60: 241-276.

[36] Murphy, K.J., "Incentives, Learning and Compensation; A Theoretical and Empirical Investigation of Managerial Labor Contracts" *The Rand Journal of Economics*, Spring 1986, vol. 17, no. 3: 59-76.

[37] Murphy, K.J., "Incentive Corporate Performance and Managerial Remuneration: An empirical Analysis" *Journal of Accounting and Economics*, 1985, vol. 7, no. 4: 11-42.

[38] Paarsh, H and Shearer, B.S., "Salaries, Piece Rates, and Intertemporal Productivity: A Study of Tree Planters in British Columbia" *Unpublished manuscript*, University of Laval, 1996.

[39] Prendergast, C.,"What Happens with Firms? A Survey of Empirical Evidence on Compensation Policies", *NBER Working paper 5802*, October 1996.

[40] Prendergast, C.,"The Provision of Incentives in Firms, *Journal of Economic Literature*, Vol. 37, no. 1, March 1999: 7-63.

[41] Petersen, T., "Reward Systems and the Distribution of Wages" *Journal of Law, Economics and Organization*, September 1991, vol. 7, 131-157.

[42] Petersen, T., "Payment Systems and the Structure of Inequality: Conceptual Issues and a Analysis of Salespersons in Department Stores" *American Journal of Sociology*, July 1992, vol. 98, 76-104.

[43] Rosenthal, R.W., and A.Weiss, "Mixed-Strategy Equilibrium in a Market with Asymmetric Information." *Review of Economic Studies*, 1994 LI: 333-342.

[44] Rothschild, M., and Stiglitz, J.,"Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, February 1976: 629-649.

[45] Stiglitz, J., "The Causes and Consequences of the Dependence of Quality on Price." *Journal of Economic Literature*, March 1987: vol. XXV 1-48.

[46] Stiglitz, J., and Weiss, A., "Credit Rationing in Markets with Imperfect Information" *American Economic Review*, June 1981, vol. 71, 393-09.

[47] Wise, A., "Human Capital vs Signalling Explanations of Wages." *The Journal of Economics Perspectives*, Fall 1995, vol. 9, no. 4: 133-54.

Appendices

A The Equilibrium Concept: Notation and Definitions

The main goal of the appendices is to state and prove formally all the propositions in the text and show that we have uniqueness of equilibrium when pure strategies are allowed.[28]

The three stage screening game is described as follows. There are two set of players, firms and workers. We denote the set of types by $\Theta = \{\overline{\theta}, \underline{\theta}\}$, with the common knowledge prior probability of $\theta = \overline{\theta}$ given by $\mu \in [0,1]$. Each firm offers a contract $C_i \in \Psi = \{C_i = (\overline{w}_i, \underline{w}_i) \in \Re : \overline{w}_i \geq \underline{w}_i \geq 0$, for $i = 1, ..., N\}$. Workers seeing the set of contracts offered $C = \cup_{i=1}^{i=N} C_i$ decides to which contract to apply. Firms seeing that, respond with either an acceptance of rejection of each contract application. A worker' pure strategy is denoted by $\sigma_\theta = (\sigma_{\theta 1, ...,} \sigma_{\theta N})$, where $\sigma_{\theta i} : \Theta \times C \rightarrow \{$not apply, apply$\}$ and a firm's pure strategy is denoted by $(\rho_i, \gamma) = (\rho_i, \gamma_{1,}..., \gamma_N)$ for all C_i, where $\rho_i : \Theta \times \Psi \rightarrow \{$reject, accept$\}$ and $\gamma_i : \Psi \rightarrow \{$not offer, offer$\}$.

Definition 1 $\Upsilon \equiv (\widehat{\mu}, \sigma_\theta, \rho, C)$ *is a pure Perfect Bayesian Equilibrium if:*

$P_1 : \forall C_i, \rho_i \in \arg\max \sum_{i=1}^{N} \sum_{\underline{\theta}}^{\overline{\theta}} \{\widehat{\mu}(\theta \mid C_i) \chi(\sigma_{\theta i}) \pi(\theta, C_i)\} \chi(\rho_i),$

where $\pi(\theta, C_i) \equiv \theta\alpha\overline{y} + (1-\theta)\alpha\underline{y} - (\theta\overline{w}_i + (1-\theta)\underline{w}_i)$, $\chi(\rho_i) = 1$ if ρ_i =accept and 0 otherwise, $\chi(\sigma_{\theta i}) = 1$ if $\sigma_{\theta i}$ =apply and 0 otherwise.

$P_2 : \forall \theta, \sigma_\theta \in \arg\max V(\theta, \Upsilon),$

where $V(\theta, \Upsilon) \equiv \sum_{i=1}^{N} (\theta U(\overline{w}_i) + (1-\theta) U(\underline{w}_i)) \chi(\sigma_{\theta i}) \chi(\rho_i)$

$P_3 : \forall C_i \in C, \mu \in [0,1], \gamma_k \in \arg\max \pi(\mu, \Upsilon),$

where $\pi(\mu, \Upsilon) \equiv \sum_{i=1}^{N} \{\sum_{\underline{\theta}}^{\overline{\theta}} \widehat{\mu}(\theta \mid C_i) \chi(\sigma_{\theta i}) \pi(\theta, C_i)\} \chi(\rho_i) \chi(\gamma_i)$, where $\chi(\gamma_{\theta i}) = 1$ if $\gamma_{\theta i}$ =offer and 0 otherwise

[28] The same result holds when mix strategies are allowed. For the sake of simplicity we focus on pure strategies. For a more formal justification of why focus only in pure strategies see Mailath (1992).

B : $\widehat{\mu}(\cdot)$ is Bayesian consistent with prior μ, firms' and workers' equilibrium strategies and observed actions whenever possible, *i.e.*, $\sum_{\underline{\theta}}^{\overline{\theta}} \widehat{\mu}(\theta \mid C_i)\chi(\sigma_{\theta i}) > 0$, otherwise $\widehat{\mu}(\cdot)$ is arbitrarily chosen. As usual the game is solved backwards, starting from stage 3 and rolling back the optimal strategies up to stage 1.

We denote the signalling sub-game by G and the set of pure strategy PBEs for the signalling sub-game by $PSE(G)$.

Definition 2 $\Lambda \equiv (\widehat{\mu}, \sigma_\theta, \rho) \in PSE(G)$ *defeats* $\Lambda' \equiv (\widehat{\mu}', \sigma'_\theta, \rho'_i) \in PSE(G)$ *if* $\exists C_i \in C$ *such that:*

C1: $\forall \theta \in \Theta : \sigma'_\theta(C_i) \neq 0$ and $K \equiv \{\theta \in \Theta : \sigma_\theta(C_i) = 1\} \neq \phi$,

C2: $\forall \theta \in K : V(\theta, \Lambda) \geq V(\theta, \Lambda')$ and $\exists \theta \in K : V(\theta, \Lambda) > V(\theta, \Lambda')$; and

C3: $\exists \theta \in K : \widehat{\mu}'(\theta \mid C_i) \neq \frac{\mu(\theta)\beta(\theta)}{\sum_{\theta' \in \Theta} \mu(\theta')\beta(\theta')} \equiv \mu(\theta, \beta(\theta))$ for any $\beta : \Theta \to [0, 1]$ satisfying

(i) $\theta' \in K$ and $V(\theta', \Lambda) > V(\theta', \Lambda')$, $\beta(\theta') = 1$ and

(ii) $\theta' \notin K$, $\beta(\theta') = 0$.

We say that a PBE Λ is undefeated if there is no other PBE Λ' that defeats Λ.

Definition 3 *The three stage screening game has an equilibrium if the set of contracts offer give rises to an undefeated PBE of the signalling sub-game;* i.e., *stages 2 and 3, with respect to all possible PBEs that may arise from any feasible set of contracts that firms may offer in stage 1.*

Let define

$$\left\{C^s_{\overline{\theta}}, C^s_{\underline{\theta}}\right\} \in \arg \max_{\{\overline{w}^s_\theta, \underline{w}^s_\theta\}^{\overline{\theta}}_{\underline{\theta}} \in C_i} \overline{\theta} U(\overline{w}^s_{\overline{\theta}}) + (1 - \overline{\theta}) U(\underline{w}^s_{\overline{\theta}}) \tag{7}$$

subject to

$$\underline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\underline{\theta}}) \geq \underline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\overline{\theta}}), \tag{$(IC\underline{\theta})$}$$

$$\overline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\overline{\theta}}) \geq \overline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\underline{\theta}}), \tag{$(IC\overline{\theta})$}$$

$$\underline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\underline{\theta}}) \geq U(0), \tag{$(IR\underline{\theta})$}$$

$$y(\overline{\theta}) \geq \overline{\theta}\overline{w}^s_{\overline{\theta}} + (1-\overline{\theta})\underline{w}^s_{\overline{\theta}}, \tag{$(ZP\overline{\theta})$}$$

$$y(\underline{\theta}) \geq \underline{\theta}\overline{w}^s_{\underline{\theta}} + (1-\underline{\theta})\underline{w}^s_{\underline{\theta}}. \tag{$(ZP\underline{\theta})$}$$

and

$$\{C^p\} \in \arg\max_{\{\overline{w}^p,\underline{w}^p\}\in C_i} \overline{\theta}U(\overline{w}^p) + (1-\overline{\theta})U(\underline{w}^p) \tag{8}$$

subject to

$$\underline{\theta}U(\overline{w}^p) + (1-\underline{\theta})U(\underline{w}^p) \geq U(0), \tag{$(IR\underline{\theta})$}$$

$$y\left(\widehat{\theta}\right) \geq \widehat{\theta}\overline{w}^p + (1-\widehat{\theta})\underline{w}^p, \tag{$(ZP\widehat{\theta})$}$$

where $\widehat{\theta} = \mu\overline{\theta} + (1-\mu)\underline{\theta}$.

Let define also the first-best contract

$$C^*_{\underline{\theta}} \in \arg\max\{V(\underline{\theta}, C_i) : \pi(\underline{\theta}, C_i) \geq 0, C_i \in C\}, \tag{9}$$

and

$$C^*_{\overline{\theta}} \in \arg\max\{V(\overline{\theta}, C_i) : \pi(\overline{\theta}, C_i) \geq 0, C_i \in C\}.$$

Our assumptions ensure that $C^*_{\underline{\theta}}$ and $C^*_{\overline{\theta}}$ exit and are unique. Moreover, $V(\theta, C^*_{\theta}) > 0$ and $\pi(\theta, C^*_{\theta}) = 0$ for $\theta \in \{\overline{\theta}, \underline{\theta}\}$.

Finally, because Preferences satisfy the Spence-Mirrless's single crossing property (hereafter, SCP) and $U(w)$ is strictly concave $\frac{d}{d\overline{w}}(\frac{V(\overline{\theta},C)}{V(\underline{\theta},C)}) > 0$; that is, paying a higher wage when the high output is realized increases high ability workers' expected utility, proportionally more than increases low ability workers' expected utility.

B Proof of Propositions 1.

The proof of proposition 1 requires to prove first a series of lemmas.

Lemma 1 *For any pairs of contracts with the same utility level, firms always prefer the one with less wage dispersion.*

Proof. For any contract $C_i \equiv (\overline{w}_i, \underline{w}_i)$ that offers a higher wage for a higher realization of output to a θ-type worker a firm's iso-profit curve is steeper than the workers indifference curve. Recall that the expected profit coming from a θ-worker is $\pi(C_i, \theta) = \theta(\alpha\overline{y} - \overline{w}_i)) + (1-\theta)(\alpha\underline{y} - \underline{w}_i))$. The absolute value of the slope of firm i's iso-profit curve when a θ-worker is hired is $\left|\frac{d\underline{w}_i}{d\overline{w}_i} |_{d\pi(C_i,\overline{\theta})=0}\right| = \dfrac{\theta}{(1-\theta)}$ and the absolute value of the slope of the indifference curve for θ-worker is $\left|\frac{d\underline{w}_i}{d\overline{w}_i} |_{dU(C_i^k,\overline{\theta})=0}\right| = \dfrac{U'(\overline{w}_i)\theta}{U'(\underline{w}_i)(1-\theta)}$. Since $\overline{w}_i \geq \underline{w}_i$ and $U(\cdot)$ is strictly concave, it follows that $U'(\underline{w}_i) \geq U'(\overline{w}_i)$, therefore the iso-profit curve is steeper than the indifference curve for any ability level, which means that firm i's expected profit increases as the contract offered moves toward the fixed-wage contract through the same indifference curve. ■

Lemma 2 *In any PBE, denoted by Υ, low-ability workers' equilibrium payoff is at least as large as the payoff that they would obtain under perfect information; that is, $V(\underline{\theta}, \Upsilon) \geq V(\underline{\theta}, C^*_{\underline{\theta}}, 1)$.*[29]

Proof. We will prove this by contradiction. Suppose not, then there exists a PBE, Υ', such that $V(\underline{\theta}, \Upsilon') < V(\underline{\theta}, C^*_{\underline{\theta}}, 1)$. By lemma 1 there is a full insurance contract, C', that provides low-ability workers with at least the same expected payoff than Υ' does, $V(\underline{\theta}, \Upsilon') = V(\underline{\theta}, C', 1)$, and yields, due to SCP, positive expected profits when is chosen by either low- or high-ability workers or both ($\pi(\theta, C') > 0, \forall \theta \in \Theta$). This implies that C' yields positive expected profits for any beliefs that firms could hold, and therefore, all firms offering C' accept all the applicants to C'. Since, we assume w.l.o.g. that there is a firm i which receive a small number of applicants, firm i can deviate and offer C''_i, where C''_i is such that $V(\underline{\theta}, C''_i, 1) > V(\underline{\theta}, C', 1)$ and attracts all workers that are applying to C'. In addition, since $\pi(\theta, C') > 0$ for any $\widehat{\mu} \in [0,1]$ it follows by continuity of the profit function that $\pi(\widehat{\mu}, C''_i) > 0$ for any $\widehat{\mu} \in [0,1]$. Therefore, no matter which are firm i's beliefs it has a profitable deviation contradicting that Υ is PBE. ■

Lemma 3 *For any PBE it is always possible to find another PBE in which along the equilibrium path at most two of the contracts offered at stage 2 are chosen,* i.e., *there are at most two $C_i \in C$ with either $\sigma_{\overline{\theta}}(C_i) > 0$ or $\sigma_{\underline{\theta}}(C_i) > 0$.*

Since we are more concerned with equilibrium outcomes than equilibrium strategies, the proof consists in finding for any PBE new strategies that yields the same equilibrium outcomes but at most two contracts are chosen in equilibrium.

Proof. Take any PBE and denote it by Υ. Suppose that at least 3 contracts offered in stage 1 are chosen, *i.e.*, either $\sigma_{\overline{\theta}}(C_i) > 0$ or $\sigma_{\underline{\theta}}(C_i) > 0$ or both for at least three contracts.

[29] With some abuse of notation $V(\underline{\theta}, \Upsilon)$ is the expected payoff that a θ-worker gets in the PBE Υ.

Let $C_{\overline{\theta}}$ be the contract, among all $C_i \in C$ such that $\sigma_{\overline{\theta}}(C_i) > 0$, that gives rises to the highest stage 3 beliefs $\widehat{\mu}(\overline{\theta} \mid C_i)$ and let define $C_{\underline{\theta}}$ in the same way, but $C_{\underline{\theta}}$ maximizes $\widehat{\mu}(\underline{\theta} \mid C_i)$, i.e., $C_{\overline{\theta}}$ $(C_{\underline{\theta}})$ is the contract most often chosen by high (low) ability workers.

Let $\underline{\mu} = \widehat{\mu}(\underline{\theta} \mid C_{\underline{\theta}})$ and $\overline{\mu} = \widehat{\mu}(\overline{\theta} \mid C_{\overline{\theta}})$ and define the following strategies;

$\sigma'_{\overline{\theta}}(C_{\overline{\theta}}) = \frac{\overline{\mu}(\mu-\underline{\mu})}{\widehat{\mu}(\overline{\mu}-\underline{\mu})}$, $\sigma'_{\overline{\theta}}(C_{\underline{\theta}}) = \frac{\underline{\mu}(\overline{\mu}-\mu)}{\widehat{\mu}(\overline{\mu}-\underline{\mu})}$, $\sigma'_{\underline{\theta}}(C_{\overline{\theta}}) = \frac{(1-\overline{\mu})(\mu-\underline{\mu})}{(1-\widehat{\mu})(\overline{\mu}-\underline{\mu})}$ and $\sigma'_{\underline{\theta}}(C_{\underline{\theta}}) = \frac{(1-\underline{\mu})(\overline{\mu}-\mu)}{(1-\widehat{\mu})(\overline{\mu}-\underline{\mu})}$.

Consider the following strategies:

Stage 1: Firms offer the same sets of contracts.

Stage 2: High ability workers' play $C_{\overline{\theta}}$ with probability $\sigma'_{\overline{\theta}}(C_{\overline{\theta}})$ and $C_{\underline{\theta}}$ with probability $1 - \sigma'_{\overline{\theta}}(C_{\overline{\theta}})$, and $\sigma_{\overline{\theta}}(C_i) = 0$ otherwise.

Low ability workers' play $C_{\overline{\theta}}$ with probability $\sigma'_{\underline{\theta}}(C_{\overline{\theta}})$ and $C_{\underline{\theta}}$ with probability $1 - \sigma'_{\underline{\theta}}(C_{\overline{\theta}})$, and $\sigma_{\underline{\theta}}(C_i) = 0$ otherwise.

Stage 3: $\rho'_i(C_i) = \rho_i(C_i)$, $\forall C_i \in C$.

It is easy to show that this new strategies give rises to same beliefs in stage 3, therefore if $\rho_i(C_i)$ was an equilibrium under Υ it must be an equilibrium under the new strategies. In stage 2, a worker of type θ is willing to randomize between two or more contracts if and only if $V(\theta, C''_i)\rho_i(C''_i) = V(\theta, C'_i)\rho_i(C'_i)$. Since, stage 3 strategies have not changed, firms are offering the same sets of contracts at stage 1, $\pi(\widehat{\mu}, C_i) = 0$, $\forall C_i \in C$ such that $\sigma'_{\theta}(C_{\theta}) > 0$, $\forall \theta \in \Theta$ and all of them give the same expected utility, then it must be case that choosing at most two contracts yields the same expected utility to a θ-worker, $\forall \theta \in \Theta$ and the same expected profits to each firm ■

Given this lemma from now on we will concentrate in the case in which each firm offer at most two contracts.[30]

Lemma 4 *In any fully separating PBE, denoted by* Υ*, low-ability workers' equilibrium payoff, is* $V(\underline{\theta}, C^*_{\underline{\theta}})$*, the payoff that they would obtain in the perfect information case.*

Proof. Let the contract chosen only by low-ability workers be denoted by $C_{\underline{\theta}}$. By lemma 2 the equilibrium payoff of this contract is so that $V(\underline{\theta}, C_{\underline{\theta}}, \rho_{\underline{\theta}}) \geq V(\underline{\theta}, C^*_{\underline{\theta}})$. By lemma 1, any contract $C_{\underline{\theta}} \neq C^s_{\underline{\theta}}$ that satisfies $V(\underline{\theta}, C_{\underline{\theta}}, \rho_{\underline{\theta}}) \geq V(\underline{\theta}, C^s_{\underline{\theta}})$ yields negative expected profits when is chosen only by low ability workers, $\pi(\underline{\theta}, C_{\underline{\theta}}) < 0$, therefore $\rho_i(C_{\underline{\theta}}) = 0$. So, $C_{\underline{\theta}}$ will be accepted with positive probability if and only if yields non-negative profits when chosen only by low ability workers. This, implies that $V(\underline{\theta}, C_{\underline{\theta}}, \rho_{\underline{\theta}}) < V(\underline{\theta}, C^s_{\underline{\theta}})$, contradicting lemma 2. Therefore, the only possible contract that is chosen only by low ability workers and is accepted in equilibrium is $C^*_{\underline{\theta}}$. This proves that $V(\underline{\theta}, C_{\underline{\theta}}, \rho_{\underline{\theta}}) = V(\underline{\theta}, C^*_{\underline{\theta}})$. ■

Lemma 5 *In any fully separating UPBE, denoted by* Λ*, high-ability workers' equilibrium payoff is at least as large as* $V(\overline{\theta}, C^s_{\overline{\theta}})$*.*

Proof. Suppose not, then there exists fully separating, Λ', such that $V(\overline{\theta}, \Lambda') < V(\overline{\theta}, C^s_{\overline{\theta}})$.

Let $C' = \{C'_1, C^*_{\underline{\theta}}, C_1\}$, where $V(\overline{\theta}, C_1) \geq V(\overline{\theta}, C^s_{\overline{\theta}})$. By lemmas 2, 3 and 4, it must be the case that $V(\overline{\theta}, C'_1, \rho'_1) > V(\overline{\theta}, C^*_{\underline{\theta}}, \rho'_{\underline{\theta}})$, $V(\underline{\theta}, C^*_{\underline{\theta}}, \rho'_{\underline{\theta}}) \geq V(\underline{\theta}, C'_1, \rho'_1)$. Hence, $\sigma'_{\overline{\theta}}(C'_1) = 1$, $\sigma'_{\underline{\theta}}(C'_1) \in [0, 1)$, $\sigma'_{\underline{\theta}}(C^*_{\underline{\theta}}) \in [0, 1]$, $\widehat{\mu}'(\underline{\theta} \mid C^*_{\underline{\theta}}) = 1$ and $\widehat{\mu}'(\overline{\theta} \mid C'_1) \in (\mu, 1]$. This is an equilibrium if and only if $V(\overline{\theta}, C_1)\rho'_1 < V(\overline{\theta}, C^s_{\overline{\theta}})$ and $V(\underline{\theta}, C^*_{\underline{\theta}}, \rho'_{\underline{\theta}}) \geq V(\underline{\theta}, C_1, \rho'_1)$, which requires that $\pi(\widehat{\mu}'(\overline{\theta} \mid C_1), C_1) = 0$. Therefore, $\widehat{\mu}'(\overline{\theta} \mid C_1) \in [0, \widehat{\mu}'^*(\overline{\theta} \mid C_1)]$, where $\widehat{\mu}'^*(\overline{\theta} \mid C_1)$ solves

[30] There is some lost of generality here, since there are some equilibria that use more than two contracts, though these equilibria have the same outcomes than the corresponding equilibria with two contracts. Given that our main concern are the equilibrium outcome and that the refinement that we use in the SSG select as unique equilibria the best separating equilibrium and the pareto optimal pooling equilibrium, this restriction has no affect on the solution of the whole game.

$\pi(\widehat{\mu}'(\overline{\theta} \mid C_1), C_1) = 0$. Notice that $\widehat{\mu}'^*(\overline{\theta} \mid C_1) < \mu$, otherwise some firms will offer C_1 and accept all the applicants to C_1 and at least break even. These firms break even since all high- and low-ability workers will apply to C_1 because $V(\overline{\theta}, C_1) \geq V(\overline{\theta}, C^s_{\overline{\theta}}) > V(\overline{\theta}, C'_1)$ and any contract such that $V(\overline{\theta}, C_1) \geq V(\overline{\theta}, C^s_{\overline{\theta}})$ satisfies the following $V(\underline{\theta}, C_1) \geq V(\underline{\theta}, C'_1)$.

Let $C = \{C_1, C^*_{\underline{\theta}}, C'_1\}$. be the set of offers in the PBE Λ. By lemmas 2, 3 and 4, it must be the case that $V(\overline{\theta}, C_1, \rho_1) > V(\overline{\theta}, C^*_{\underline{\theta}}, \rho^*_{\underline{\theta}})$, $V(\overline{\theta}, C_1, \rho_1) \geq V(\overline{\theta}, C^s_{\overline{\theta}})$, $V(\underline{\theta}, C^*_{\underline{\theta}}, \rho^*_{\underline{\theta}}) \geq V(\underline{\theta}, C_1, \rho_1)$. Hence, $\sigma_{\overline{\theta}}(C_1) = 1$, $\sigma_{\underline{\theta}}(C_1) \in [0, 1]$, $\sigma_{\underline{\theta}}(C^*_{\underline{\theta}}) \in (0, 1]$, $\widehat{\mu}(\underline{\theta} \mid C^*_{\underline{\theta}}) = 1$ and $\widehat{\mu}(\overline{\theta} \mid C_1) \in (\mu, 1]$.

Since in Λ', $\sigma_\theta(C_1) = 0$, $\forall \theta \in \Theta$, and in Λ, $\sigma_{\overline{\theta}}(C_1) = 1$, $K = \{\overline{\theta}\}$ (if $\sigma_{\underline{\theta}}(C_1) > 0$, then $K = \Theta$), which is non-empty, satisfying C1 of our refinement concept. Condition C2 is satisfied because $V(\underline{\theta}, \Lambda) = V(\underline{\theta}, \Lambda')$ and $V(\overline{\theta}, \Lambda) > V(\overline{\theta}, \Lambda')$. In the case in which $K = \{\overline{\theta}\}$ condition C3 imposes that $\beta(\overline{\theta}) = 1$, $\beta(\underline{\theta}) = 0$, therefore $\mu(\overline{\theta}, \beta(\overline{\theta})) = 1$ which is different from $\widehat{\mu}'(\overline{\theta} \mid C_1)$. Therefore, Λ defeats any separating PBE, Λ', in which $V(\overline{\theta}, \Lambda') < V(\overline{\theta}, C^s_{\overline{\theta}})$. If $K = \Theta$ condition C3 imposes that $\beta(\overline{\theta}) = 1$, $\beta(\underline{\theta}) \in [0, 1]$, therefore $\mu(\overline{\theta}, \beta(\overline{\theta})) \in [\mu, 1]$ and $\mu(\underline{\theta}, \beta(\underline{\theta})) \in [0, \mu]$, which are different from $\widehat{\mu}'(\overline{\theta} \mid C_1)$ and $\widehat{\mu}'(\underline{\theta} \mid C_1)$ because of $\widehat{\mu}'^*(\overline{\theta} \mid C_1) < \mu$.

This proves that any fully separating UPBE that offers to high-ability workers an expected payoff lower than $V(\overline{\theta}, C^s_{\overline{\theta}})$ is defeated by a PBE that offers at least $V(\overline{\theta}, C^s_{\overline{\theta}})$ to high-ability workers and $V(\underline{\theta}, C^*_{\underline{\theta}})$ to low-ability workers. ■

Lemma 6 *In any fully separating, denoted by Λ, high-ability workers' equilibrium payoff is equal to the expected payoff from contract $C^s_{\overline{\theta}}$, $V(\overline{\theta}, C^s_{\overline{\theta}})$.*

Proof. Suppose there exist a PBE, denoted by $\Lambda^{'}$, so that $V(\overline{\theta}, \Lambda^{'}) \neq V(\overline{\theta}, C^s_{\overline{\theta}})$. By lemmas 5 and 4, in any separating, $\forall C_i \in C$ such that $\sigma_{\underline{\theta}}(C_i) > 0$, $V(\underline{\theta}, C_i, \rho_i) = V(\underline{\theta}, C^*_{\underline{\theta}})$ and $\forall C_i \in C$ such that $\sigma_{\overline{\theta}}(C_i) > 0$, $V(\overline{\theta}, C_i, \rho_i) \geq V(\overline{\theta}, C^s_{\overline{\theta}})$ and $\pi(\overline{\theta}, C_i) \geq 0$. Moreover, since $\pi(\overline{\theta}, C^s_{\overline{\theta}}) = 0$ and the iso-profit function is negatively sloped any contract $C_i \neq C^s_{\overline{\theta}}$ such that $V(\overline{\theta}, C_i, \rho_i) > V(\overline{\theta}, C^s_{\overline{\theta}})$ and $\pi(\overline{\theta}, C_i) \geq 0$ must satisfy the following $U(\overline{w}_i) < U(\overline{w}^s)$ and

$U(\underline{w}_i) > U(\underline{w}^s)$. By SCP any contract C_i such that $V(\overline{\theta}, C_i, \rho) > V(\overline{\theta}, C^s_{\overline{\theta}})$ is also preferred by low-ability workers, so it must be the case that $\frac{V(\underline{\theta}, C^s_{\underline{\theta}})}{V(\underline{\theta}, C_i)} \geq \rho(C_i)$ and $\rho(C_i) > \frac{V(\overline{\theta}, C^s_{\overline{\theta}})}{V(\overline{\theta}, C_i)}$, $\forall C_i \in C$ such that $\sigma_{\overline{\theta}}(C_i) = 1$ and $\sigma_{\underline{\theta}}(C_i) > 0$. Combining these two inequalities we get that $\frac{V(\overline{\theta}, C_i)}{V(\underline{\theta}, C_i)} > \frac{V(\overline{\theta}, C^s_{\overline{\theta}})}{V(\underline{\theta}, C^s_{\underline{\theta}})} = \frac{V(\overline{\theta}, C^s_{\overline{\theta}})}{V(\underline{\theta}, C^s_{\overline{\theta}})}$, where the equality follows from the definition of $C^s_{\overline{\theta}}$ and $C^*_{\underline{\theta}}$. This plus $U(\overline{w}_i) < U(\overline{w}^s)$ and $U(\underline{w}_i) > U(\underline{w}^s)$ contradicts $\frac{d}{d\overline{w}}(\frac{V(\overline{\theta}, C)}{V(\underline{\theta}, C)}) > 0$. Therefore, there is no fully separating or semi-separating PBE where $V(\overline{\theta}, \Lambda') \neq V(\overline{\theta}, C^s_{\overline{\theta}})$. ■

Lemma 7 *The contract C^p is a PBE of the signalling sub-game.*

Proof. Recall that C^p is the solution to program II above.

The first-order conditions for this program are,

$$\overline{\theta}U'(\overline{w}^p) + \widehat{\theta}\lambda_1 - \lambda_2\underline{\theta}U'(\overline{w}^p) = 0, \tag{10}$$

$$(1-\overline{\theta})U'(\underline{w}^p) + (1-\widehat{\theta})\lambda_1 - \lambda_2(1-\underline{\theta})U'(\underline{w}^p) = 0. \tag{11}$$

If we assume that low-ability workers' participation constraint is not binding, then $\lambda_2 = 0$ and from 10 and 11 it follows that,

$$\frac{U'(\overline{w}^p)}{U'(\underline{w}^p)} = \frac{\widehat{\theta}}{\overline{\overline{\theta}}}\frac{(1-\overline{\theta})}{(1-\widehat{\theta})}. \tag{12}$$

Hence, we have that $\overline{w}^p_i > \underline{w}^p_i \geq 0$. This proves that low-ability workers' participation constraint is not binding, therefore, assuming $\lambda_2 = 0$ impose no restrictions on the solution.

To prove that there is a PBE of the SSG that sustain C^p as a PBE, notice first that by definition C^p maximizes high-ability workers' expected payoff and breaks even only at the population average probability of success, therefore $\pi(\underline{\theta}, C^p) < 0$. To prove that C^p can be supported as PBE consider the following strategies:

Stage 1: $C_i = \{C^p\}$ for $i \leqq k < N$ and $C_i = \{C'\}$ for $i > k$.

Stage 2: $\sigma_{\overline{\theta}}(C^p) = 1$ and $\sigma_{\underline{\theta}}(C^p) = 1$.

Stage 3: $\rho_i(C^p) = 1$ and $\rho_i(C') = 1$, $\forall C' \in C$ such that $\pi(\underline{\theta}, C') \geq 0$.

On-the-equilibrium-path beliefs: $\widehat{\mu}(\overline{\theta} \mid C^p) = \mu$.

Off-the-equilibrium-path beliefs: $\widehat{\mu}(\overline{\theta} \mid C') = 0$, $\forall C' \neq C^p$.

It is easy to check that these strategies satisfy the PBE requirements. ■

Lemma 8 *The contracts $C^s_{\underline{\theta}}$ and $C^s_{\overline{\theta}}$ are a PBE of the signalling sub-game.*

Proof. Recall that contract $C^s_{\overline{\theta}}$ maximizes high-ability workers expected payoff subject to that low-ability workers do not mimic high-ability workers.

The first-order conditions for program II when high-ability workers' incentive compatibility constraint is ignored are

$$
\begin{aligned}
\overline{w}^s_{\underline{\theta}}: &\quad \lambda_2 \underline{\theta} + \lambda_1 \underline{\theta} U'(\overline{w}^s_{\underline{\theta}}) \leqq 0, \\
\underline{w}^s_{\underline{\theta}}: &\quad \lambda_2 (1 - \underline{\theta}) + \lambda_1 (1 - \underline{\theta}) U'(\underline{w}^s_{\underline{\theta}}) \leqq 0, \\
\overline{w}^s_{\overline{\theta}}: &\quad (\overline{\theta} - \lambda_1 \underline{\theta}) U'(\overline{w}^s_{\overline{\theta}}) + \overline{\theta} \lambda_3 \leqq 0, \\
\underline{w}^s_{\overline{\theta}}: &\quad (1 - \overline{\theta} - \lambda_1 (1 - \underline{\theta})) U'(\underline{w}^s_{\overline{\theta}}) + (1 - \overline{\theta}) \lambda_3 \leqq 0.
\end{aligned}
$$

From this it follows immediately that $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}} = w^s_{\underline{\theta}}$ and from $ZP\underline{\theta}$ we know that $w^s_{\underline{\theta}} = y(\underline{\theta})$.

Solving for λ_1, λ_2 and λ_3,

$$
\lambda_1 = \tfrac{1}{M^s} \left((1 - \overline{\theta}) \overline{\theta} (U'(\overline{w}^s_{\overline{\theta}}) - U'(\underline{w}^s_{\overline{\theta}})) \right) > 0 \tag{13}
$$

$$
\lambda_2 = \tfrac{1}{M^s} \left((1 - \overline{\theta}) \overline{\theta} (U'(\overline{w}^s_{\overline{\theta}}) - U'(\underline{w}^s_{\overline{\theta}})) \right) U'(w^s_{\underline{\theta}}) > 0 \tag{14}
$$

$$\lambda_3 = -\frac{(\overline{\theta}-\underline{\theta})}{M^s} U'(\overline{w}^s_{\overline{\theta}}) U'(\underline{w}^s_{\overline{\theta}}) > 0 \tag{15}$$

where $M^s = (1-\overline{\theta})\underline{\theta}U'(\overline{w}^s_{\overline{\theta}}) - \overline{\theta}(1-\underline{\theta})U'(\underline{w}^s_{\overline{\theta}}) < 0$.

Checking that high-ability workers' incentive compatibility constraint is not binding is simple, just notice that since $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ and $\overline{\theta} > \underline{\theta}$, $\overline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\overline{\theta}}) > \underline{\theta}U(\overline{w}^s_{\overline{\theta}}) + (1-\underline{\theta})U(\underline{w}^s_{\overline{\theta}}) = U(w^s_{\underline{\theta}}) = \overline{\theta}U(\overline{w}^s_{\underline{\theta}}) + (1-\overline{\theta})U(\underline{w}^s_{\underline{\theta}})$.

So, $C^s_{\overline{\theta}}$ gives the higher expected payoff to high-ability workers among all separating. In order to sustain the solution to program II as a PBE consider the following strategies and beliefs.

Stage 1: $C_i = \{C^s_{\underline{\theta}}\}$ for $i \leqq k < N$ and $C_i = \{C^s_{\overline{\theta}}\}$ for $i > k$.

Stage 2: $\sigma_{\overline{\theta}}(C^s_{\overline{\theta}}) = 1$ and $\sigma_{\underline{\theta}}(C^s_{\underline{\theta}}) = 1$.

Stage 3: $\rho_i(C^s_{\overline{\theta}}) = 1$ and $\rho_i(C^s_{\underline{\theta}}) = 1$.

On-the-equilibrium-path beliefs: $\widehat{\mu}(\overline{\theta} \mid C^s_{\overline{\theta}}) = 1$ and $\widehat{\mu}(\underline{\theta} \mid C^s_{\underline{\theta}}) = 1$.

Off-the-equilibrium-path beliefs: $\widehat{\mu}(\overline{\theta} \mid C') = 0$, $\forall C' \neq C^s_{\overline{\theta}}$.

It is easy to check that these strategies satisfy the PBE requirements. ■

Proof of Proposition 1 For $\mu < \widetilde{\mu}$, there is a unique UPBE in which every firm offers $C = \{C^s_{\underline{\theta}}, C^s_{\overline{\theta}}\}$, low-ability workers choose $C^s_{\underline{\theta}}$ with probability 1 and high ability workers choose $C^s_{\overline{\theta}}$ with probability 1 and every firm accepts $C^s_{\underline{\theta}}$ and $C^s_{\overline{\theta}}$ with probability 1. For $\mu \geq \widetilde{\mu}$, there is a unique UPBE in which every firm offers $C_i = \{C^p\}$ $\forall i$, both types of workers choose C^p with probability 1 and C^p is accepted with probability 1.

Proof.

Part 1: $\mu \leqq \widetilde{\mu}$.

Suppose there exits a PBE, Λ, that defeats Λ^s[31]; that is, there exists a contract $C_j \in C$ such that $\sigma^s_\theta(C_j) = 0$, $\forall\theta \in \Theta$ and $\sigma_\theta(C_j) > 0$ for some $\theta \in \Theta$ so that $C2$ and $C3$ are satisfied.

[31] Ω^s denotes the PBE of SSG that sustain the contracts $C^s_{\overline{\theta}}$ and $C^s_{\underline{\theta}}$ as a PBE.

By lemmas 4 and 5 in any fully separating PBE, $V(\overline{\theta},\Lambda) = V(\overline{\theta},C^s_{\overline{\theta}})$ and $V(\underline{\theta},\Lambda) = V(\underline{\theta},C^s_{\underline{\theta}})$. This implies that there is no fully separating PBE different from Λ^s that satisfies conditions $C2$ and $C3$. Therefore, there is no separating PBE that defeats Λ^s. By definition of $\widetilde{\mu}$, when $\mu < \widetilde{\mu}$, the highest payoff that high-ability workers can get in a pooling PBE is such that $V(\overline{\theta},C^p) < V(\overline{\theta},C^s_{\overline{\theta}})$. Since C^p will be a pooling equilibrium it is the case that $C_j \in C$ such that $\sigma^s_{\theta}(C^p) = 0, \forall\theta \in \Theta$ and $\sigma_{\theta}(C^p) > 0$, therefore, $K = \Theta$ and condition $C2$ is immediately violated. Therefore, when $\mu \leqq \widetilde{\mu}$ there is no PBE that defeats Λ^s. Uniqueness follows from the fact if a semi-separating equilibrium exists only low-ability workers play mix strategies in which case, the only contract that breaks even when chosen only by high-ability workers is $C^s_{\overline{\theta}}$. Furthermore, any contract chosen by low-ability workers must promise a payoff of $V(\underline{\theta},C^s_{\underline{\theta}})$ which is equal to $V(\underline{\theta},C^s_{\overline{\theta}})$.

Part 2: $\mu > \widetilde{\mu}$.

When $\mu > \widetilde{\mu}$, by definition of $\widetilde{\mu}$ the highest payoff that high-ability workers can get in a pooling PBE is such that $V(\overline{\theta},C^p) > V(\overline{\theta},C^s_{\overline{\theta}})$ and $V(\underline{\theta},C^p) > V(\underline{\theta},C^s_{\underline{\theta}})$. Therefore, it is trivial to show that Λ^s is defeated by Λ^p[32]. The question is whether there is another pooling equilibrium besides Λ^p that is undefeated.

Suppose there exists a PBE denoted by $\Lambda^{p'}$ that defeats Λ^p. Recall that by definition, C^p is, among all possible contracts used in any PBE, the contract that yields the highest payoff to high-ability workers.

Let $C' = \{C^p, C^{p'}\}$ and $\sigma'_{\theta}(C^p) = 0, \forall\theta \in \Theta$ and $C = \{C^p, C^{p'}\}$ and $\sigma_{\theta}(C^{p'}) = 0, \forall\theta \in \Theta$. In this case $K = \Theta \neq \phi$, then $C2$ fails since any pooling PBE $\Lambda^{p'}$ different from Λ^p is such that $V(\overline{\theta},\Lambda^{p'}) < V(\overline{\theta},\Lambda^p)$. ■

[32] We denote by Ω^p the PBE that sustains C^p as a PBE.

B.1 Proof of proposition 2.

The first-order conditions for program EPS in section 4.2 when the high-ability worker's participation constraint is ignored are

$$\overline{w}^s_{\underline{\theta}}: \quad \left(\lambda_1\underline{\theta} - \lambda_2\overline{\theta} + \lambda_3\underline{\theta}\right) U'(\overline{w}^s_{\underline{\theta}}) - \lambda_5\underline{\theta} = 0, \tag{16}$$

$$\underline{w}^s_{\underline{\theta}}: \quad \left(\lambda_1(1-\underline{\theta}) - \lambda_2\left(1-\overline{\theta}\right) + \lambda_3(1-\underline{\theta})\right) U'(\underline{w}^s_{\underline{\theta}}) - \lambda_5(1-\underline{\theta}) = 0, \tag{17}$$

$$\overline{w}^s_{\overline{\theta}}: \quad (\overline{\theta} + \lambda_2\overline{\theta} - \lambda_1\underline{\theta})U'(\overline{w}^s_{\overline{\theta}}) - \overline{\theta}\lambda_4 = 0, \tag{18}$$

$$\underline{w}^s_{\overline{\theta}}: \quad (\left(1-\overline{\theta}\right) + \lambda_2\left(1-\overline{\theta}\right) - \lambda_1(1-\underline{\theta}))U'(\underline{w}^s_{\overline{\theta}}) - (1-\overline{\theta})\lambda_4 = 0, \tag{19}$$

$$e^s_{\overline{\theta}} \quad -1 + \lambda_1 - \lambda_2 + \lambda_4 y_e\left(e^s_{\overline{\theta}}, \theta\right) = 0, \tag{20}$$

$$e^s_{\underline{\theta}} \quad -\lambda_1 + \lambda_2 - \lambda_3 + \lambda_5 y_e\left(e^s_{\underline{\theta}}, \theta\right) = 0. \tag{21}$$

It follows from 20 that $1 + \lambda_2 = \lambda_1 + \lambda_4 y_e\left(e^s_{\overline{\theta}}, \theta\right)$. Hence, either $\lambda_1 > 0$ or $\lambda_4 > 0$ or both. It follows from 18, 19 and 20 that

$$\frac{\lambda_1}{\lambda_4} = -\frac{\overline{\theta}\left(U'(\underline{w}^s_{\overline{\theta}})y_e\left(e^s_{\overline{\theta}}, \overline{\theta}\right) - 1\right)}{\triangle\theta U'(\overline{w}^s_{\overline{\theta}})} \geq 0. \tag{22}$$

Hence, $U'(\underline{w}^s_{\overline{\theta}})y_e\left(e^s_{\overline{\theta}}, \theta\right) - 1 \leqq 0$. If we plug 22 into 19, we obtain that

$$y_e\left(e^s_{\overline{\theta}}, \overline{\theta}\right) = \left(1-\overline{\theta}\right)\frac{1}{U'(\underline{w}^s_{\overline{\theta}})} + \overline{\theta}\frac{1}{U'(\overline{w}^s_{\overline{\theta}})}. \tag{23}$$

It follows from 21, 16 and 17 that $\lambda_1 + \lambda_3 = \lambda_2 + \lambda_5 y_e\left(e^s_{\underline{\theta}}, \theta\right)$ and

$$\frac{\lambda_2}{\lambda_5} = \frac{\underline{\theta}\left(U'(\underline{w}^s_{\underline{\theta}})y_e\left(e^s_{\underline{\theta}}, \underline{\theta}\right) - 1\right)}{\triangle\theta U'(\overline{w}^s_{\underline{\theta}})} \geq 0. \tag{24}$$

Hence, $U'(\underline{w}^s_{\underline{\theta}})y_e\left(e^s_{\underline{\theta}}, \underline{\theta}\right) - 1 \geq 0$. If we plug 24 into 17, we obtain that

$$y_e\left(e^s_{\underline{\theta}}, \underline{\theta}\right) = (1-\underline{\theta})\frac{1}{U'(\overline{w}^s_{\underline{\theta}})} + \underline{\theta}\frac{1}{U'(\underline{w}^s_{\underline{\theta}})}. \tag{25}$$

It follows from 24, 25 and $\overline{w}^s_{\underline{\theta}} \geq \underline{w}^s_{\underline{\theta}}$ that $\lambda_2 = 0$, while it follows from 22, 23, and $\overline{w}^s_{\overline{\theta}} \geq \underline{w}^s_{\overline{\theta}}$ that $\lambda_1 > 0$. Because $\lambda_2 = 0$ and $\lambda_1 > 0$, we have that $\lambda_5 > 0$.

Because $\lambda_2 = 0$, it follows from 16 and 17 that $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}} = w^s_{\underline{\theta}}$ and from $ZP\underline{\theta}$ we know that $w^s_{\underline{\theta}} = y\left(e^s_{\underline{\theta}}, \underline{\theta}\right)$. This coupled with 25 and the assumption that $U\left(y\left(e^*_{\underline{\theta}}, \underline{\theta}\right)\right) - e^*_{\underline{\theta}} > 0$ implies that $e^s_{\underline{\theta}} = e^*_{\underline{\theta}}$ and that $\lambda_3 = 0$.

Using 18, 19 and the fact that $\lambda_2 = 0$. It can be shown that

$$\lambda_1 = \tfrac{1}{M^s}\left((1-\overline{\theta})\overline{\theta}(U'(\overline{w}^s_{\overline{\theta}}) - U'(\underline{w}^s_{\overline{\theta}}))\right), \tag{26}$$

$$\lambda_4 = -\tfrac{(\overline{\theta}-\underline{\theta})}{M^s}U'(\overline{w}^s_{\overline{\theta}})U'(\underline{w}^s_{\overline{\theta}}), \tag{27}$$

where $M^s = (1-\overline{\theta})\underline{\theta}U'(\overline{w}^s_{\overline{\theta}}) - \overline{\theta}(1-\underline{\theta})U'(\underline{w}^s_{\overline{\theta}}) < 0$.

Hence, $\lambda_4 > 0$ and $\lambda_1 > 0$. It also follows from this that the optimal contract $C^s_{\overline{\theta}}$ is such that $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ and $\overline{\theta}\overline{w}^s_{\overline{\theta}} + (1-\overline{\theta})\underline{w}^s_{\overline{\theta}} = y_e\left(e^s_{\underline{\theta}}, \theta\right)$.

Furthermore, it follows from 23 that if $h(w) \equiv \frac{1}{U'(w)}$ is concave $e^s_{\overline{\theta}} \leqq e^*_{\overline{\theta}}$, while if $h(w)$ is convex $e^s_{\overline{\theta}} > e^*_{\overline{\theta}}$.

Finally notice that $(IC\underline{\theta})$, $(IC\overline{\theta})$ and $(IR\underline{\theta})$ implies $(IR\overline{\theta})$, constrain that was ignored at the time of solving the problem above.

Similarly, in a pooling equilibrium firms maximize high-ability workers' expected payoff subject to low-ability workers' participation constraint, $(IR\underline{\theta})$, and a zero expected profit constraint when evaluated at the population average probability of success, $\widehat{\theta}$. The first'order conditions for program EPP in section 4.2 when a high-ability worker's participation constraint is ignored since it is implied by the low-ability worker's participation constraint are

$$\overline{\theta}U'(\overline{w}^p) - \widehat{\theta}\lambda_1 + \lambda_2\underline{\theta}U'(\overline{w}^p) = 0, \tag{28}$$

$$(1-\overline{\theta})U'(\underline{w}^p) - (1-\widehat{\theta})\lambda_1 + \lambda_2(1-\underline{\theta})U'(\underline{w}^p) = 0. \tag{29}$$

$$-1 + \lambda_1 y_e\left(e^p, \widehat{\theta}\right) - \lambda_2 = 0 \tag{30}$$

It follows from 30 that $\lambda_1 > 0$ and that $\lambda_1 = \frac{1+\lambda_2}{y_e\left(e^p,\widehat{\theta}\right)}$, hence $y\left(e^p,\widehat{\theta}\right) = \widehat{\theta}\overline{w}^p + (1-\widehat{\theta})\underline{w}^p$. It also follows from the equilibrium concept that a pooling equilibrium exists if and only if both, high- and low-ability worker are at least as well-off as in the separating equilibrium. This implies that if a pooling equilibrium exists $\underline{\theta}U(\overline{w}^p)+(1-\underline{\theta})U(\underline{w}^p)-e^p > U\left(y\left(e^*_{\underline{\theta}},\underline{\theta}\right)\right)-e^*_{\underline{\theta}} > 0$. Hence, it can be assumed that $\lambda_2 = 0$. Hence, from 28, 29 and 30 it follows that,

$$U^{'}(\overline{w}^p) = \frac{1}{y_e\left(e^p,\widehat{\theta}\right)}\frac{\widehat{\theta}}{\overline{\theta}}, \tag{31}$$

$$U^{'}(\underline{w}^p) = \frac{1}{y_e\left(e^p,\widehat{\theta}\right)}\frac{(1-\widehat{\theta})}{(1-\overline{\theta})}, \tag{32}$$

and

$$y_e\left(e^p,\widehat{\theta}\right) = \widehat{\theta}h(\overline{w}^p) + (1-\widehat{\theta})h(\underline{w}^p),$$

where $h(w) \equiv \frac{1}{U^{'}(w)}$.

The first two equations above imply that $\overline{w}^p > \underline{w}^p$ because $\frac{\widehat{\theta}}{\overline{\theta}} < \frac{(1-\widehat{\theta})}{(1-\overline{\theta})}$. The third equation implies that $e^p < e^*_{\underline{\theta}}$ and $e^*_{\overline{\theta}} \gtreqless e^p$ if h is concave, and $e^p > e^*_{\overline{\theta}}$ and $e^*_{\underline{\theta}} \gtreqless e^p$ if h is convex.

B.2 Proof of proposition 3.

In this case a firms' problem when the firms sort workers out is given the following program

$$\left\{C^s_{\overline{\theta}}, C^s_{\underline{\theta}}\right\} \in \arg \max_{\{\overline{w}^s_\theta,\underline{w}^s_\theta,p_\theta\}^{\overline{\theta}}_{\underline{\theta}}\in C_i} \overline{\theta}\,(p_{\overline{\theta}})\,U(\overline{w}^s_{\overline{\theta}}) + (1-\overline{\theta}\,(p_{\overline{\theta}}))U(\underline{w}^s_{\overline{\theta}}) \tag{33}$$

subject to

$$\underline{\theta}\,(p_{\underline{\theta}})\,U(\overline{w}^s_{\underline{\theta}}) + (1-\underline{\theta}\,(p_{\underline{\theta}}))U(\underline{w}^s_{\underline{\theta}}) \geq \underline{\theta}\,(p_{\overline{\theta}})\,U(\overline{w}^s_{\overline{\theta}}) + (1-\underline{\theta}\,(p_{\overline{\theta}}))U(\underline{w}^s_{\overline{\theta}}), \tag{$IC\underline{\theta}$}$$

$$\overline{\theta}\left(p_{\overline{\theta}}\right)U(\overline{w}^s_{\overline{\theta}})+(1-\overline{\theta}\left(p_{\overline{\theta}}\right))U(\underline{w}^s_{\overline{\theta}})\geq\overline{\theta}\left(p_{\underline{\theta}}\right)U(\overline{w}^s_{\underline{\theta}})+(1-\overline{\theta}\left(p_{\underline{\theta}}\right))U(\underline{w}^s_{\underline{\theta}}), \tag{$IC\overline{\theta}$}$$

$$\underline{\theta}\left(p_{\underline{\theta}}\right)U(\overline{w}^s_{\underline{\theta}})+(1-\underline{\theta}\left(p_{\underline{\theta}}\right))U(\underline{w}^s_{\underline{\theta}})\geq 0, \tag{$IR\underline{\theta}$}$$

$$\theta\left(p_{\theta}\right)\overline{w}^s_{\theta}+(1-\theta\left(p_{\theta}\right))\underline{w}^s_{\theta}\leqq y(\theta)-\gamma C\left(p_{\theta}\right)\ \text{ for }\theta\in\left\{\underline{\theta},\overline{\theta}\right\}, \tag{$ZP\theta$}$$

while when firms pool the workers under the same contract is

$$\{C^p\}\in\arg\max_{\{\overline{w}^p,\underline{w}^p,p^p\}\in C_i}\overline{\theta}\left(p\right)U(\overline{w}^p)+(1-\overline{\theta}\left(p\right))U(\underline{w}^p) \tag{34}$$

subject to

$$\underline{\theta}\left(p\right)U(\overline{w}^p)+(1-\underline{\theta}\left(p\right))U(\underline{w}^p)\geq 0, \tag{$IR\underline{\theta}$}$$

$$\widehat{\theta}\left(p\right)\overline{w}^p+(1-\widehat{\theta}\left(p\right))\underline{w}^p\leqq y\left(\widehat{\theta}\right)-\gamma C\left(p\right). \tag{$ZP\widehat{\theta}$}$$

It follows immediately that the optimal contract offered when the equilibrium is pooling and firms can monitor ability satisfies the same condition than when firms cannot monitor ability.

The first-order conditions for the case in which the equilibrium is separating when $IC\overline{\theta}$ and $IR\underline{\theta}$ are ignored are given by

$$\begin{aligned}
\overline{w}^s_{\underline{\theta}}:&\quad -\lambda_2\underline{\theta}\left(p_{\underline{\theta}}\right))+\lambda_1\underline{\theta}\left(p_{\underline{\theta}}\right))U'(\overline{w}^s_{\underline{\theta}})\leqq 0,\\
\underline{w}^s_{\underline{\theta}}:&\quad -\lambda_2(1-\underline{\theta}\left(p_{\underline{\theta}}\right))+\lambda_1(1-\underline{\theta}\left(p_{\underline{\theta}}\right))U'(\underline{w}^s_{\underline{\theta}})\leqq 0,\\
p^s_{\underline{\theta}}:&\quad \lambda_1\left(2\underline{\theta}-1\right)\left(U(\overline{w}^s_{\underline{\theta}})-U(\underline{w}^s_{\underline{\theta}})\right)-\lambda_2\left(\left(2\underline{\theta}-1\right)\left(\overline{w}^s_{\underline{\theta}}-\underline{w}^s_{\underline{\theta}}\right)+\gamma C'\left(p^s_{\underline{\theta}}\right)\right)\lesseqqgtr 0,\\
\overline{w}^s_{\overline{\theta}}:&\quad (\overline{\theta}\left(p_{\overline{\theta}}\right)-\lambda_1\underline{\theta}\left(p_{\overline{\theta}}\right))U'(\overline{w}^s_{\overline{\theta}})-\overline{\theta}\left(p_{\overline{\theta}}\right)\lambda_3\leqq 0,\\
\underline{w}^s_{\overline{\theta}}:&\quad \left(1-\overline{\theta}\left(p_{\overline{\theta}}\right)-\lambda_1(1-\underline{\theta}\left(p_{\overline{\theta}}\right))\right)U'\left(\underline{w}^s_{\overline{\theta}}\right)-(1-\overline{\theta}\left(p_{\overline{\theta}}\right))\lambda_3\leqq 0,\\
p^s_{\overline{\theta}}:&\quad \left(\left(2\overline{\theta}-1\right)-\lambda_1\left(2\underline{\theta}-1\right)\right)\left(U(\overline{w}^s_{\overline{\theta}})-U(\underline{w}^s_{\overline{\theta}})\right)-\lambda_3\left(\left(2\overline{\theta}-1\right)\left(\overline{w}^s_{\overline{\theta}}-\underline{w}^s_{\overline{\theta}}\right)+\gamma C'\left(p^s_{\overline{\theta}}\right)\right)\lesseqqgtr 0.
\end{aligned}$$

It readily follows from the first and second equation that $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}}$, and using this in the third equation it follows that $p^s_{\underline{\theta}} = \frac{1}{2}$.

It readily follows from the 4th and 5th equation that for all $p^s_{\overline{\theta}} > \frac{1}{2}$, $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ and that

$$\lambda_1 = \tfrac{1}{M^s}(1 - \overline{\theta}\,(p_{\overline{\theta}}))\overline{\theta}\,(p_{\overline{\theta}})\,(U'(\overline{w}^s_{\overline{\theta}}) - U'(\underline{w}^s_{\overline{\theta}})) > 0, \tag{35}$$

$$\lambda_2 = \tfrac{1}{M^s}(1 - \overline{\theta}\,(p_{\overline{\theta}}))\overline{\theta}\,(p_{\overline{\theta}})\,(U'(\overline{w}^s_{\overline{\theta}}) - U'(\underline{w}^s_{\overline{\theta}}))U'(w^s_{\underline{\theta}}) > 0, \tag{36}$$

$$\lambda_3 = -\tfrac{(\overline{\theta}-\underline{\theta})}{M^s}U'(\overline{w}^s_{\overline{\theta}})U'(\underline{w}^s_{\overline{\theta}}) > 0, \tag{37}$$

where $M^s = (1 - \overline{\theta}\,(p_{\overline{\theta}}))\underline{\theta}\,(p_{\overline{\theta}})\,U'(\overline{w}^s_{\overline{\theta}}) - \overline{\theta}\,(p_{\overline{\theta}})\,(1 - \underline{\theta}\,(p_{\overline{\theta}}))U'(\underline{w}^s_{\overline{\theta}}) < 0$.

Lemma 9 $p^s_{\overline{\theta}} > \frac{1}{2}$.

Suppose not, then $\overline{\theta}\,(p_{\overline{\theta}}) = \underline{\theta}\,(p_{\overline{\theta}}) = \frac{1}{2}$. It follows from the $IR\underline{\theta}$ that the maximum wage that a firm is willimg to pay to a $\underline{\theta}$-worker is $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}} = y(\underline{\theta})$ and therefore, a $\underline{\theta}$-worker's expected utility is $U\,(y(\underline{\theta}))$. Because of $IC\underline{\theta}$, $U\,(y(\underline{\theta})) \geq \frac{1}{2}U(\overline{w}^s_{\overline{\theta}}) + \frac{1}{2}U(\underline{w}^s_{\overline{\theta}})$. By concavity of U this implies that $y(\underline{\theta}) \geq \frac{1}{2}\overline{w}^s_{\overline{\theta}} + \frac{1}{2}\underline{w}^s_{\overline{\theta}}$; that is, $\frac{1}{2}\overline{w}^s_{\overline{\theta}} + \frac{1}{2}\underline{w}^s_{\overline{\theta}} < y\left(\overline{\theta}\right)$ and therefore firms make positive profits. But this implies by continuity that there is a firm, the one that attract the least number of high-ability workers, that can deviate by choosing a monitoring precision $\frac{1}{2} + \varepsilon$, and wages $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ such that $U\,(y(\underline{\theta})) \geq \underline{\theta}\left(\frac{1}{2} + \varepsilon\right)U(\overline{w}^s_{\overline{\theta}}) + \left(1 - \underline{\theta}\left(\frac{1}{2} + \varepsilon\right)\right)U(\underline{w}^s_{\overline{\theta}})$ and $\overline{\theta}\left(\frac{1}{2} + \varepsilon\right)U(\overline{w}^s_{\overline{\theta}}) + \left(1 - \overline{\theta}\left(\frac{1}{2} + \varepsilon\right)\right) > U\,(y(\underline{\theta})) \geq \frac{1}{2}U(\overline{w}^s_{\overline{\theta}}) + \frac{1}{2}U(\underline{w}^s_{\overline{\theta}})$ and therefore attract all high-ability workers, no low-ability workers and have larger profits. Hence, in a separating equilibrium it must be true that $p^s_{\overline{\theta}} > \frac{1}{2}$, $\overline{w}^s_{\overline{\theta}} > \underline{w}^s_{\overline{\theta}}$ and , $\overline{w}^s_{\underline{\theta}} = \underline{w}^s_{\underline{\theta}} = y(\underline{\theta})$.

The first-order conditions for the case in which the equilibrium is pooling when the $IR\underline{\theta}$

is ignored are given by

$$\begin{aligned}
\overline{w}^p &: \quad -\lambda_1 \widehat{\theta}\left(p^p\right) + \overline{\theta}\left(p^p\right) U'(\overline{w}^p) \leqq 0, \\
\underline{w}^p &: \quad -\lambda_1 (1 - \widehat{\theta}\left(p^p\right)) + (1 - \overline{\theta}\left(p^p\right)) U'(\overline{w}^p) \leqq 0, \\
p^p &: \quad \left(2\overline{\theta} - 1\right)\left(U(\overline{w}^p) - U(\underline{w}^p)\right) - \lambda_1 \left(\left(2\widehat{\theta} - 1\right)\left(\overline{w}^p - \underline{w}^p\right) + \gamma C'\left(p^p\right)\right) \lesseqqgtr 0.
\end{aligned}$$

It readily follows from the first two equations that when $p^p > \frac{1}{2}$, $\overline{w}^p > \underline{w}^p$. Furthermore, it readily follows from either the first or second equation that $\lambda_1 > 0$, so $\widehat{\theta}\left(p^p\right)\overline{w}^p + (1 - \widehat{\theta}\left(p^p\right))\underline{w}^p = y\left(\widehat{\theta}\right) - \gamma C\left(p^p\right)$. While if $p^p = \frac{1}{2}$, $\widehat{\theta}\left(p^p\right) = \overline{\theta}\left(p^p\right)$, so $\overline{w}^p = \underline{w}^p = y\left(\widehat{\theta}\right)$.

Finally, it follows from the third equation that

$$p_p = \begin{cases} \frac{1}{2} & \text{if } -\lambda_1 \gamma C'\left(\frac{1}{2}\right) < 0, \\ [\frac{1}{2}, 1) & \text{if } \left(2\overline{\theta} - 1\right)\left(U(\overline{w}^p) - U(\underline{w}^p)\right) - \lambda_1 \left(\left(2\widehat{\theta} - 1\right)\left(\overline{w}^p - \underline{w}^p\right) + \gamma C'\left(p^p\right)\right) = 0, \\ 1 & \text{if } \left(2\overline{\theta} - 1\right)\left(U(\overline{w}^p) - U(\underline{w}^p)\right) - \lambda_1 \left(\left(2\widehat{\theta} - 1\right)\left(\overline{w}^p - \underline{w}^p\right) + \gamma C'\left(1\right)\right) > 0. \end{cases} \tag{38}$$

Solving for λ_1 from either the first or the second equation and using the concavity of U, it is easy to show that $\frac{(2\overline{\theta}-1)}{\lambda_1}\left(U(\overline{w}^p) - U(\underline{w}^p)\right) - \left(2\widehat{\theta} - 1\right)\left(\overline{w}^p - \underline{w}^p\right) \geq 0$ for all $p \geq \frac{1}{2}$. Thus, if γ is sufficiently large, then the optimal monitoring intensity is $p^p = \frac{1}{2}$. That is, it is optimal to pay offer a degenerate menu containing the straight-salary contract that promises $\overline{w}^p = \underline{w}^p = y\left(\widehat{\theta}\right)$.